

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds  
(without alignments)  
11177.792 Million cell updates/sec

Title: US-10-069-056-4  
Perfect score: 2019  
Sequence: 1 agcgtggaatgcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2019	100.0	2019	6	AX137739 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	V01115 Minute viru
4	2017.4	99.9	5149	14	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137743 Sequence
6	2015.8	99.8	2019	6	AX137747 Sequence
7	2015.8	99.8	2019	6	AX137751 Sequence
8	1905.4	94.2	5085	14	M12032 Minute viru
9	1902.2	94.2	4764	14	M0U34253
10	1902.2	94.2	5087	14	X02481 Mouse parvo
11	1899	94.1	5144	14	U12469 Mouse parvo
12	1873	92.8	4764	14	M0U34254
13	1862.2	92.2	4773	14	H0U34255
14	1855.8	91.9	4761	14	MVU34256
15	1771	87.7	5121	6	CQ786765 Sequence
16	1662.2	82.3	4904	14	AF321230
17	1659	82.2	4795	14	AF332884
18	1657.4	82.1	4927	14	KRU079033
19	1656.6	82.1	5135	14	PVRSEQ

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARH1	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	52.0	5124	14	PVPVC	M38246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCY1A	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAFPVNS1	X55115 Feline panl
40	1043.6	51.7	5124	14	PVCPVC	M38245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVNRFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137739	Sequence 4 from Patent EP1077260.	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137739					
DEFINITION	Sequence 4 from Patent EP1077260.					
ACCESSION	AX137739					
VERSION	AX137739.1	GI:14273913				
KEYWORDS						
SOURCE						
ORGANISM						

Mice minute virus  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
Nuesch, J. and Rommelaere, J.  
Parvovirus ns1 variants  
Patent: EP 1077260-A 4 21-FEB-2001;  
Deutsches Krebsforschungszentrum' Stiftung des Oeffentlichen Rechts  
(DE)

FEATURES

source  
Location/Qualifiers  
1..2019  
/organism="Mice minute virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10794"  
/note="Parvovirus NS1 variant"  
1..2019  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC39991.1"  
/db\_xref="GI:14273914"  
/translation="MAGNAYSDEVLGATWNLKEKSNQEVFSVFVKNNVQLNGDKIGW  
NSYKELQDELSLQORGAETWDSQEDMEWETTVDEMTKKQVIFDSLVKKCLFEVL  
NTKNIFPGVNNFVQHEWGDKQCHVLIGGKDSQAQGWRRQLNVVWSRWLVTA  
CNVQTPAERIKLREIAEDNWTLLTYKQTKDYKTRCVLFGNNMAYYFLTKKGIS  
TSPPRDGGYFLSDSGKTNFLKEGHELVSKLYTDDMPETVETVTTVAQETKRGRI  
OTKEVAIKTTLKELVHKVTSPEDMMWQPSYIEMWAQPGGENLKNLTLEICLTLL  
ARTKAPDLILEKETSCLTNESLPDTRICRFAGHNVYKVCHACCVLNFQGGKR  
NTVLFHGPASTGKSIIFAQIAQAGNVGNCYNAANVPNDCTNKNLIWEEAGNFGQ  
QNQKFAICSGGTIRIDQKGSQKLEPTVIMTINENITVVRIGCEERPEPTQIR  
RMLNHLHTLFDGLVDKNWPMICAWLVKNGYQSTWASVYCAKRGKVPDWSENWAE  
PKVPTPINLLGSARSPTTPKSTPLSQNTALTPLASDLEDLALPEPSTNTVPVAGTAE  
TQNTGTBAGSKAQDQGLSPTWSIEIEDLRACFGAEPLKKDFSEPLND"

CDS

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0; Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCTGMAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTAAAGCAAAA	60
DB	1	ATGGCTGMAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTAAAGCAAAA	60
QY	61	AGTAACCAAGGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
DB	61	AGTAACCAAGGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
QY	121	GATATCGATGGAATAGTTTACAAAAAGAGCTCGAGGAGGACGAGCTGAATCTTTTACA	180
DB	121	GATATCGATGGAATAGTTTACAAAAAGAGCTCGAGGAGGAGGAGCTGAATCTTTTACA	180
QY	181	CGAGGACGGAAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAACCAACAGTGGAT	240
DB	181	CGAGGACGGAAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAACCAACAGTGGAT	240
QY	241	GAATGACCAAAAGCAAGTATTCAATTTTGGTTTAAAAAATGTTTATTGAA	300
DB	241	GAATGACCAAAAGCAAGTATTCAATTTTGGTTTAAAAAATGTTTATTGAA	300
QY	301	GTGCTTAAACAAAGAAATATATTTCCTGTGTGATGTTAAATGTTTGTGCAACATGAATGG	360
DB	301	GTGCTTAAACAAAGAAATATATTTCCTGTGTGATGTTAAATGTTTGTGCAACATGAATGG	360
QY	361	GGAAAAACCAAGCTGGCACTGCCATGTAATAATTGGAGGAAGGACTTTAGTCAAGCT	420
DB	361	GGAAAAACCAAGCTGGCACTGCCATGTAATAATTGGAGGAAGGACTTTAGTCAAGCT	420
QY	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGTAAACAGCC	480
DB	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGTAAACAGCC	480
QY	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540
DB	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540
QY	541	GAGTGGGTACTCTACTTATAGCATAGCAAAACCAAAAAGACTATACCAAGTGT	600
DB	541	GAGTGGGTACTCTACTTATAGCATAGCAAAACCAAAAAGACTATACCAAGTGT	600
QY	601	GTTCTTTTTGGAAACATGATGTGTTACTATTTTTTAACTAAAAAGAAATTAAGCACTAGT	660
DB	601	GTTCTTTTTGGAAACATGATGTGTTACTATTTTTTAACTAAAAAGAAATTAAGCACTAGT	660
QY	661	CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAACCTAACTTTTTTA	720
DB	661	CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAACCTAACTTTTTTA	720
QY	721	AAAGAGGGGAGCGGCATCTAGTGAGCAAACTATACATGATGATCGCGCCAGAAACG	780
DB	721	AAAGAGGGGAGCGGCATCTAGTGAGCAAACTATACATGATGATCGCGCCAGAAACG	780
QY	781	GTTGAAACCAACATGATTAACCTGCGAGGAAACTAAGCGCGCAGAAATCTAACTAAAAA	840
DB	781	GTTGAAACCAACATGATTAACCTGCGAGGAAACTAAGCGCGCAGAAATCTAACTAAAAA	840
QY	841	GAAGTTGCTATTAAAACTACATTAAGAGCTGGTGCAATAAAGAGTAACCTCACAGAG	900
DB	841	GAAGTTGCTATTAAAACTACATTAAGAGCTGGTGCAATAAAGAGTAACCTCACAGAG	900
QY	901	GACTGATGATGATGAGCGGACAGTGTACATTTGAATGATGGCTCAACNAGTGGAGAA	960
DB	901	GACTGATGATGATGAGCGGACAGTGTACATTTGAATGATGGCTCAACNAGTGGAGAA	960
QY	961	AACCTGCTGAAAAATACGCTAGAGATTGTTACACTAACTTAGCCAGAACCAAAACAGCA	1020
DB	961	AACCTGCTGAAAAATACGCTAGAGATTGTTACACTAACTTAGCCAGAACCAAAACAGCA	1020
QY	1021	TTTGACTTAATTTTGAAGAAAGCTGAAACCAAGCAAACTTAACCACTTTCACTGCCTGAC	1080



KEYWORDS	Mice minute virus	Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAAGCAAT	540
SOURCE	Mice minute virus	Qy	541	GAGTGGGTACTCTACTTACTTATAGCATATAGCAAAACCAAAAGACTATACCAAGTGT	600
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	Db	541	GAGTGGGTACTCTACTTACTTATAGCATATAGCAAAACCAAAAGACTATACCAAGTGT	600
REFERENCE	1 Nueesch, J. and Rommelaere, J.	Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
AUTHORS	Parvovirus nel variants	Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
TITLE	Patent: EP 1077260-A 1 21-FEB-2001;	Qy	661	CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	720
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)	Db	661	CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	720
FEATURES	Location/Qualifiers	Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
source	1. .2019	Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
CDS	/organism="Mice minute virus"	Qy	781	GTGTAACACACAGTAACCACTGCGCAGGAAACTTAAGCGGCGAGAAATTAACCTAAAAAA	840
	/mol_type="unassigned DNA"	Db	781	GTGTAACACACAGTAACCACTGCGCAGGAAACTTAAGCGGCGAGAAATTAACCTAAAAAA	840
	/db_xref="taxon:10794"	Qy	841	GAAAGTTCCTATTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACCTCACCAGAG	900
	/note="Wildtype Parvovirus NS1"	Db	841	GAAAGTTCCTATTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACCTCACCAGAG	900
	/notes="unnamed protein product"	Qy	901	GACTGATGATGATGAGCAGACAGTATACATTGAAATGATGCTCAACAGGTGGAGAA	960
	/codon_start=1	Db	901	GACTGATGATGATGAGCAGACAGTATACATTGAAATGATGCTCAACAGGTGGAGAA	960
	/protein_id="CAC39989.1"	Qy	961	AACTGCTGAAATAACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA	1020
	/translation="MAGNAYSDVLGATNWLKEKSNQVFSFVFKNNVOLNGKIDIGW	Db	961	AACTGCTGAAATAACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA	1020
	NSYKLEQDELKSLQGEATTWQSEDMETWEDTKKQVFIIDSLVKCLFPEVL	Qy	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
	NTKNIIPGDVNWFOHEWKGDKQCHVLIGKDFSOAGCKWRRLNLYWWSRLVTA	Db	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
	CNVQLTPAERIKLREAEADNEWTLITKHKQTKDYTKCVLPGNMIAYFLTKKIS	Qy	1081	ACAAGAACCTGCGAGAAATTTTGGCTGCGGCAAGCAAGCAAGCAAGCAAGCAAGCA	1140
	TSPPRGYFLSDSGWKNTFLKEGERHLVSKYITDDMPETVEITTTTAQSTKGR	Db	1081	ACAAGAACCTGCGAGAAATTTTGGCTGCGGCAAGCAAGCAAGCAAGCAAGCAAGCA	1140
	QTKKEYSIKTLLKELVHKRTPSPEDMMQPDYSIEMQAOPGGENLLKNTLEICTL	Qy	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA	1200
	ARTKPAFLILEAKETSKLTFSLPDRICRIFAPGWNVYKCHACVCLNKGGR	Db	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA	1200
	NVLFPASPASTGKSIATAIAQAVGNVCYNAAVNFDFNDCTNKNLIWVEEAGNFGQ	Qy	1201	GCAGACAGGCAAACTCTATTATTGCACAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
	OVNQFAICSGOTIRIDOKGSKQIEPTPIVMTNENITVVRIGCEERPEHTQPIRD	Db	1201	GCAGACAGGCAAACTCTATTATTGCACAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
	RLNLIHLTHLPGDFGLVDKNPIMICAMLVKNGYQSTWASICYAKNGKTPDPNENWAE	Qy	1261	TGCTATAATGCGACCAATGTAAACCTTTTCCATTAAATGACTGTATACCAACAGAACTTGATT	1320
	KVPTPINLIGSASFPTPKSPILSONYALTPLASDLALEPSTPTNPVAGTAE	Db	1261	TGCTATAATGCGACCAATGTAAACCTTTTCCATTAAATGACTGTATACCAACAGAACTTGATT	1320
	TQNTGAGSKACQDGLSPTWSEIEBDRACFGABPLKDDFSEPLND	Qy	1321	TGGGTAGAGAAAGCTGGTAACTTTTGACAGCAAGTAACCCAGTTTAAAGCCATTGCTCT	1380
ORIGIN	Query Match 99.9%; Score 2017.4; DB 6; Length 2019;	Db	1321	TGGGTAGAGAAAGCTGGTAACTTTTGACAGCAAGTAACCCAGTTTAAAGCCATTGCTCT	1380
	Best Local Similarity 100.0%; Pred. No. 0;	Qy	1381	GGTCAAACTATTTCGCAATGATCAAAAGGAAAGGAGCAAGCAAGCAAGCAAGCAAGCA	1440
	Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Db	1381	GGTCAAACTATTTCGCAATGATCAAAAGGAAAGGAGCAAGCAAGCAAGCAAGCAAGCA	1440
		Qy	1441	GTCAATCATGACCAAAATGAGAACATTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
	1 ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAAA	Db	1501	GAAACACTCAACCAATTCAGAGACAGAAATGCTTAACTTCATCTACACATACCTTGCTCT	1560
	60	Qy	1501	GAAACACTCAACCAATTCAGAGACAGAAATGCTTAACTTCATCTACACATACCTTGCTCT	1560
	1 ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAAA	Db	1561	GGTGACTTTGGTTGGTTGGTTCACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGAAG	1620
	60	Qy	1561	GGTGACTTTGGTTGGTTGGTTCACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGAAG	1620
	61 AGTAACAGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAAA	Db			
	120	Qy			
	61 AGTAACAGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAAA	Db			
	120	Qy			
	121 GATATCGGATGGAATAGTTTACAAAGAGCTGCGAGGACGAGCTGAAATCTTTACAA	Db			
	180	Qy			
	121 GATATCGGATGGAATAGTTTACAAAGAGCTGCGAGGACGAGCTGAAATCTTTACAA	Db			
	180	Qy			
	181 CGAGGAGCGAACTACTTGGGACCAAGCGAGACATGGAATGGGAAACACAGTGGAT	Db			
	240	Qy			
	181 CGAGGAGCGAACTACTTGGGACCAAGCGAGACATGGAATGGGAAACACAGTGGAT	Db			
	240	Qy			
	241 GAAATGACCAAAAGCAAGTATTCAATTTTGAATCTTTGGTTAAAAAATGTTTATTGAA	Db			
	300	Qy			
	241 GAAATGACCAAAAGCAAGTATTCAATTTTGAATCTTTGGTTAAAAAATGTTTATTGAA	Db			
	300	Qy			
	301 GTGCTTAAACCAAGAAATATATTCCTGGTGATGTTAAATGGTTGTGCAACATGATGG	Db			
	360	Qy			
	301 GTGCTTAAACCAAGAAATATATTCCTGGTGATGTTAAATGGTTGTGCAACATGATGG	Db			
	360	Qy			
	361 GGAAGAGACCAAGCTGCGACTGCGATGATTAATTTGGAGGAAAGGACTTTAGTCAAGCT	Db			
	420	Qy			
	361 GGAAGAGACCAAGCTGCGACTGCGATGATTAATTTGGAGGAAAGGACTTTAGTCAAGCT	Db			
	420	Qy			
	421 CAAGGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	Db			
	480	Qy			
	421 CAAGGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	Db			
	480	Qy			
	481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT	Db			
	540	Qy			



|||||  
561 GTGCTTAACCAAGAATAATTTCCGGTGATGTTAAATGGTTGTGCAACATGAATGG 620  
|||||  
361 GGAAGACCAAGGCTGGCACTGTCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT 420  
|||||  
621 GGAAGACCAAGGCTGGCACTGTCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT 680  
|||||  
421 CAAGGGAATGTGGAGGAAGCACTAAATTTTACCTGGAGCAGATGGTTGGTAAACAGCC 480  
|||||  
681 CAAGGGAATGTGGAGGAAGCACTAAATTTTACCTGGAGCAGATGGTTGGTAAACAGCC 740  
|||||  
481 TGTAACTGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
|||||  
741 TGTAACTGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 800  
|||||  
541 GAGTGGGTACTCTACTTACTTATTAAGCATAAGCAACCAAAAAGAGACTATACCAAGTGT 600  
|||||  
801 GAGTGGGTACTCTACTTACTTATTAAGCATAAGCAACCAAAAAGAGACTATACCAAGTGT 860  
|||||  
601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAGAAAATAAGCACTAGT 660  
|||||  
861 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAGAAAATAAGCACTAGT 920  
|||||  
661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAGCACTTAACTTTTA 720  
|||||  
921 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAGCACTTAACTTTTA 980  
|||||  
721 AAAGAAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 780  
|||||  
981 AAAGAAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 1040  
|||||  
781 GTTGAACCAACAGTAACCACTGGCGAGGAACTAAGCGCGGCAAGTAATCAAAATAAAAAA 840  
|||||  
1041 GTTGAACCAACAGTAACCACTGGCGAGGAACTAAGCGCGGCAAGTAATCAAAATAAAAAA 1100  
|||||  
841 GAAGTGTCTATTAAACTACATTAAGAGCTGGTGATTAAGAGTAACCTCACCAGAG 900  
|||||  
1101 GAAGTGTCTATTAAACTACATTAAGAGCTGGTGATTAAGAGTAACCTCACCAGAG 1160  
|||||  
901 GACTGGATGATGATGCAGCCAGACAGTTACATTTGAATGATGGCTCAACAGAGTGGAGAA 960  
|||||  
1161 GACTGGATGATGATGCAGCCAGACAGTTACATTTGAATGATGGCTCAACAGAGTGGAGAA 1220  
|||||  
961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAACCAAAACAGCA 1020  
|||||  
1221 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAACCAAAACAGCA 1280  
|||||  
1021 TTTGACTTAATTTAGAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080  
|||||  
1281 TTTGACTTAATTTAGAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1340  
|||||  
1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAATCTATGTTAAAGTTTGCCTGCT 1140  
|||||  
1341 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAATCTATGTTAAAGTTTGCCTGCT 1400  
|||||  
1141 ATTTGCTGTGTTTAAACAGCAAGGAGGCAAAAGAAATACTGTTTATTTATGAGACCA 1200  
|||||  
1401 ATTTGCTGTGTTTAAACAGCAAGGAGGCAAAAGAAATACTGTTTATTTATGAGACCA 1460  
|||||  
1201 GCAGACAGGCAAACTCTATTATGCAACGCCATAGCAACGAGTGTGGCAATGTTGGT 1260  
|||||  
1461 GCCAGCAGGCAAACTCTATTATGCAACGCCATAGCAACGAGTGTGGCAATGTTGGT 1520  
|||||  
1261 TGCTATAATCAGCAATGTAACTTTTCCATTTTAACTGACGTACCAACAGAACTTGATT 1320  
|||||  
1521 TGCTATAATCAGCAATGTAACTTTTCCATTTTAACTGACGTACCAACAGAACTTGATT 1580  
|||||  
1321 TGGGTAGAAGAGCTGGTAACCTTTTGGACAGCAAGTAAACAGATTTAAAGCCATTGCTCT 1380  
|||||  
1581 TGGGTAGAAGAGCTGGTAACCTTTTGGACAGCAAGTAAACAGATTTAAAGCCATTGCTCT 1640  
|||||  
1381 GGTCAAACTATTGCAATGATCAAAAAGGAAAGGAGGAGCAAAACAGATTTGAACCAACCA 1440  
|||||

Db 1641 GGTCAAACTATTTCGATTGATCAAAAAGGAAAGGCAAGCAAGATTTGAACCAACCA 1700  
Qy 1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGT CAGAAATAGCTGCGAAGAAAGACCA 1500  
Db 1701 GTCATCATGACCAAAATGAGAACATTTACAGTGGT CAGAAATAGCTGCGAAGAAAGACCA 1760  
Qy 1501 GAACACACTCAACCAATCAGAGACAGANTGCTTTAACTTCATCTACACACATACCTTGCT 1560  
Db 1761 GAACACACTCAACCAATCAGAGACAGANTGCTTTAACTTCATCTACACACATACCTTGCT 1820  
Qy 1561 GGTGACTTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620  
Db 1821 GGTGACTTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1880  
Qy 1621 AATGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680  
Db 1881 AATGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1940  
Qy 1681 TCAGAAAACCTGGGGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGTTTGGGCACGC 1740  
Db 1941 TCAGAAAACCTGGGGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGTTTGGGCACGC 2000  
Qy 1741 TCACCAATTCACGACACCGAAAGCTAGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 2001 TCACCAATTCACGACACCGAAAGCTAGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060  
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG 1860  
Db 2061 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG 2120  
Qy 1861 GGCATCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920  
Db 2121 GGCATCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 2180  
Qy 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGCTGGGGAACCG 1980  
Db 2181 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGCTGGGGAACCG 2240  
Qy 1981 TTGAAGAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db 2241 TTGAAGAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2279

## RESULT 4

LOCUS MWMPG 5149 bp ss-DNA linear VRL 22-MAY-1995  
DEFINITION Minute virus of mice, complete genome.  
ACCESSION J02275 M12520 M12521 M14704  
VERSION J02275.1 GI:332293  
KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.  
SOURCE Mice minute virus  
ORGANISM Mice minute virus  
REFERENCE 1 (bases 1 to 5149)  
AUTHORS Astell,C.R., Thomson,M., Merchinsky,M. and Ward,D.C.  
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus  
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)  
MEDLINE 83143341  
PUBMED 6298737  
REFERENCE 2 (bases 1 to 5149)  
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.  
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain  
J. Virol. 57 (2), 656-669 (1986)  
MEDLINE 86115415  
PUBMED 3502703  
REFERENCE 3 (sites)  
AUTHORS Morgan,W.R. and Ward,D.C.  
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs

J. Virol. 60 (3), 1170-1174 (1986) 87061199 3783817 PUBMED COMMENT	J. Virol. 60 (3), 1170-1174 (1986) 87061199 3783817 PUBMED COMMENT	Original source text: Minute virus of mice (strain MVM(p)), passed in mouse l (variant A-9) cells. The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group. The sequence below corresponds to the plus (+) strand, also referred to as the C-strand. The minus (-) strand is also referred to as the V-strand. The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1]. The left and right halves of the genome encode two distinct, but overlapping transcriptional units. The transcripts can be summarized [1] (1 map unit (mu) = 51 bp): R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu R3 is the major transcript. There are two major open reading frames, both on the plus (or C) strand. The left side ORF (261-2279) probably encodes a non-capsid protein of 85 kd; the right side ORF probably encodes the viral capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known. revises [1]. 4870 a-65bp-a in [2]; aa in [1] [2] [3] sites; splice sites. location/Qualifiers 1. 5149 /organism="Mice minute virus" /mol_type="genomic DNA" /strain="MVM(p)" /db_xref="taxon:10794" /lab_host="mouse l (variant A-9) cell" 114. .2279 /gene="NS1" 114. .2279 /gene="NS1" /notes="putative" /codon_start=1 /product="nonstructural protein" /protein_id="AAA67108.1" /db_xref="GI:825481" /translation="MISGSGSLNQAKRWKPKVYKLLKSVYLLFFHSVSRDAQE SNOLTWAGNAYSDEVLGATNLWKEKSNQVFSFVKENVLQNGKDIGNSVYKKELOE DELKSLQRGNETTWDQSEDMEWETTTVDMTKKQVFIPLSVKKCLPEVLNKNIPFGD VNMVFEHQMGQCHGVHIGGKDFSGAQKRWRLNYSWRLVLTACNVOLTPAE RIKLRIADENNEWTLTYKHQKDYTKCVLFGNMIAYFLTKKISTSPRDGGY FLSSDSGKTHLVSKLYDDMRPETVTTVAQETKGRITQRTKKEVSIK TTLKELVHKRTVSPEDMMQPSYIEMMAQPGENLLKNTLEICTLTARTKTFADL ILEKARTSKLTNLSLPTDTRCIFAFGHNYVVKVCHAI CCVLRNQGKENTVLPHPGA STGKSLIAQAQAVNGVCYNAANVFPNDCTNKNLIWVEAGNFGQGVQVQKAI C SGQTRIDQKGGKQLEIPTPTVIMTNENITVVRIGCEERPEHTQPIRDMNLHLTH TLFGDFGLDKNWPMICAWLVKNGYQSTWASYCAKWKVPDWSNENAEKPKVPTPLN LGSARFPFTPKSTPLSQNYALPLASDLELALPEWSTPTNPVAGTAEPTQNTGEGAS KACQDGLSPTWSEI EEDLRACFGABPLKKDFSEPLNLD" 200. .>2279 /gene="NS1" 261. .2279 /gene="NS1" /codon_start=1 /product="nonstructural protein" /protein_id="AAA67109.1" /db_xref="GI:312295" /translation="MAGNAYSDEVLGATNLWKEKSNQVFSFVKENVLQNGKDIGNW	NSYKKELOEDELKSLQRGNETTWDQSEDMEWETTTVDMTKKQVFIPLSVKKCLPEVL NTKNIIPFGDVMNFMVHGWCKDQCHVILGGKDFSGAQKRWRLNYSWRLVLTAC CNVQLTPAESRIKLRIADENNEWTLTYKHQKDYTKCVLFGNMIAYFLTKKIS TSPRDGGYFLSSDSGKTHLVSKLYDDMRPETVTTVAQETKGRITQRTKKEVSIK QRTKKEVSIKTLKELVHKRTVSPEDMMQPSYIEMMAQPGENLLKNTLEICTLT ARTKTFADLILEKARTSKLTNLSLPTDTRCIFAFGHNYVVKVCHAI CCVLRNQGK NTVLPFGPASTGKSLIAQAQAVNGVCYNAANVFPNDCTNKNLIWVEAGNFGQ QVNOFKAISGOTIRIDQKGGKQLEIPTPTVIMTNENITVVRIGCEERPEHTQPIR RMLNHLTHLTLFGDFGLDKNWPMICAWLVKNGYQSTWASYCAKWKVPDWSNENAE KPKVPTPLNLSGARSFPFTPKSTPLSQNYALPLASDLELALPEWSTPTNPVAGTAE TQNTGAGSKACQDGLSPTWSEI EEDLRACFGABPLKKDFSEPLNLD" 2002. .2398 /gene="VP" 2002. .2280 /gene="VP" /note="major transcription start site" /number=1 2006. .2280 /gene="VP" /note="minor transcription start site" /number=1 2009. .2280 /gene="VP" /note="minor transcription start site" /number=1 2081. .2398 /gene="VP" /note="alternative intron" 2281. .2376 /gene="VP" /note="alternative intron" Join(2286. .2316,2399. .4557) /gene="VP1" Join(2286. .2316,2399. .4557) /gene="VP1" /codon_start=1 /protein_id="AAA67111.1" /db_xref="GI:332294" /translation="MAPPAKRAKRWPPGYKYLPGNSLDQGEPTNPDSAAAKHDE AYDOYKSKNPPLYESAADOFIDOTKADMGKGVGHYFERTKRAFPKLTATSEP GTSGVSRAGKTRTPPAYIFINOARAKKLTSSAAQSSQTSMDGTSQDPSGNVHISA RVERAADPGSGGGGGGGVGVSTYDNTQHTHYRFLGDDGWEITATLRLVHLNMP KSENYCRIRVHNTDTSYKGNMAKDDAHEQIWPMSLDANAWGVLQSPDWQYICNT MSQLNLVSLDQIEFNVLKTVTEQDLGGQAIKIYNNDLTACMVAVDSNNILPYTPAA NSMETLGFVPWKPTIASPYRYFCVDRDLSTYENQEGVHEHNVMTGPKMNSQPFIT ENTQOITLLRTGDEPATGYTFTDTSVKLTHTWQTNOLGQPLLSTPEADTDAFTL TAQSGRHGTTQMGVNWVSEAIRTPAQVGFQCPNDPEASRAGPAAAPKVPADITQGV DKRANGSVRYSTYKQHGHNWASHGHPAPERYTWDETSFGSDTKDGFQISAPLVVPPP LNGILITNANPIGTKNDIHFSNVFNSYGLPTAPSHSPVYPOQIWDKELDLHKKPLRH ITAPFVCKNNAFGQMLVRLGNLTDQYDPNGATLSRIVTYGTFFWKGLTWRAKLRA TTWNPVYOVSABEDNGNSYMTKWLPTATGNNQSVPLITRTPVARNTY" 2286. .2354 /gene="VP" /note="ORF1; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67110.1" /db_xref="GI:825482" /translation="MAPPAKRAKRGKGLRDGLVGY" 2286. .2316 /gene="VP" /number=1 2317. .2398 /gene="VP" /note="VP intron (alt.)" 2332. .2361 /gene="VP" /note="ORF3; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67112.1" /db_xref="GI:825483" /translation="MVGWGWGINV"	gene exon exon exon intron intron gene CDS CDS exon intron CDS
--	--	---	--	---

CDS	2354.. 2398 /gene="VP" /note="ORF2; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67113.1" /db_xref="GI:825484" /translation="MFNYLFPYRPEITWF" 2399..>4557 /gene="VP1" /number=2 2794.. 4557 /gene="VP1" /note="VP2" /codon_start=1 /protein_id="AAA67114.1" /db_xref="GI:332296" /translation="MSDGTQSPGSNAVHSAARVERAADPGSGSGGGGGVGVST GSVDNTHYRFLGDGWEIETALATRLVHLNMPKSENYCRIRVHNTTDTSVKGNMAKDD AHEQIWPMSLDANAGVWLQSDQWYICNTMSQLNLVSLDQEIENFVLKTVTEQDL GGQAIKIYNDLTACVMVAVDNNILPYTPAANSMETLGFYPMKPTIAGPYRYFCDV RDLSTVYENQEGVEHVMGTPKGMNSQFFTIENQOITLLRTGDFEATGTYFFDTNS VKLTHWQTNROLQPPPLSTPPEADTDAGTLTAQGRHGTQMGVNMVSEAIRTPA QVGFCDPHNDFEASRAGPFAAPKVPADITQGVDEKANGSVRSYGHKNGHENWASHGPA PERYTWDETSFGSRDPTKGFQISAPLAVPPPLNGILTNANPIGTAKNDIHFNVFNSY GPLTAFSPSPVYPOQIWDKELDLHKPLRLHTATPFVCKNNAPGQMLVRLGNLTDQ	Db	861	GTTCTTTTGGAAAACATGATTGCTTACTATTTTAACTAAAGAAAATAAGCACTAGT	920
		QY	661	CCACCAAGAGACGAGGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
		Db	921	CCACCAAGAGACGAGGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	980
exon		QY	721	AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGGCAGAAACG	780
		Db	981	AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGGCAGAAACG	1040
CDS		QY	781	GTTGAAAACCACTAGTAACCACTGCGCAGGAAAACTAAGCGCGGCAGAAATTCAAATAAAAA	840
		Db	1041	GTTGAAAACCACTAGTAACCACTGCGCAGGAAAACTAAGCGCGGCAGAAATTCAAATAAAAA	1100
		QY	841	GAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGATATAAAAGAGTAACCTCACAGAG	900
		Db	1101	GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGATATAAAAGAGTAACCTCACAGAG	1160
		QY	901	GACTGATGATGATGCGAGCCAGACAGTTACATTTGAAATGATGCTCAACACGAGTGGAGAA	960
		Db	1161	GACTGATGATGATGCGAGCCAGACAGTTACATTTGAAATGATGCTCAACACGAGTGGAGAA	1220
		QY	961	AACCTGCTGAAAAATAACGCTAGAGATTTGTACACTAACTCTTAGCCAGAACCAAAAAACG	1020
		Db	1221	AACCTGCTGAAAAATAACGCTAGAGATTTGTACACTAACTCTTAGCCAGAACCAAAAAACG	1280
		QY	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
		Db	1281	TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1340
		QY	1081	ACAAGAACCTGCGAAATTTTGTGCTTTTCATGGCTGGAACTATGTTTAAAGTTTGCCTGCT	1140
		Db	1341	ACAAGAACCTGCGAAATTTTGTGCTTTTCATGGCTGGAACTATGTTTAAAGTTTGCCTGCT	1400
		QY	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTCTGTTTATTTTCAATGAGCA	1200
		Db	1401	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTCTGTTTATTTTCAATGAGCA	1460
		QY	1201	GCAGACAGGCAAACTATATTATTGCAAGCCATAGCAAGCAGTGGCAATGTTGGT	1260
		Db	1461	GCAGACAGGCAAACTATATTATTGCAAGCCATAGCAAGCAGTGGCAATGTTGGT	1520
		QY	1261	TGCTATAATGCGAGCAATGTAAACTTTCCATTTTAATGACTGTACCAACCAAGAACTTGAT	1320
		Db	1521	TGCTATAATGCGAGCAATGTAAACTTTCCATTTTAATGACTGTACCAACCAAGAACTTGAT	1580
		QY	1321	TGGGTAGAGAGAGCTGTAACCTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380
		Db	1581	TGGGTAGAGAGAGCTGTAACCTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1640
		QY	1381	GGTCAAACTATTGCGCAATGATCAAAAAGGAAAAGGCGAGCAAAAGAGTTGAACCAACCA	1440
		Db	1641	GGTCAAACTATTGCGCAATGATCAAAAAGGAAAAGGCGAGCAAAAGAGTTGAACCAACCA	1700
		QY	1441	GTCAATCATGACCAAAATGAGAACATTTACNGTGGTCAAGATAGGCTGCGAAGAACCA	1500
		Db	1701	GTCAATCATGACCAAAATGAGAACATTTACNGTGGTCAAGATAGGCTGCGAAGAACCA	1760
		QY	1501	GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTGCCT	1560
		Db	1761	GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTGCCT	1820
		QY	1561	GGTGACTTTGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG	1620
		Db	1821	GGTGACTTTGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG	1880
		QY	1621	AATGGTTTCAATCTACCTAGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG	1680
		Db	1881	AATGGTTTCAATCTACCTAGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG	1940
		QY	1681	TCAGAAAACCTGGCGGAGCCAAAGGTGCCAACTCTCTATAAAATTTTACTAGTTCGGCAGCG	1740
		Db	1941	TCAGAAAACCTGGCGGAGCCAAAGGTGCCAACTCTCTATAAAATTTTACTAGTTCGGCAGCG	2000

```
QY 1741 TCACATTTCAGCACCGAAAGTACGCCTCTCAGCCAGAACATATGCACTAACTCCACTT 1800
Db 2001 TCACATTTCAGCACCGAAAGTACGCCTCTCAGCCAGAACATATGCACTAACTCCACTT 2060
QY 1801 GCATCGATCTGAGGACCTGGCTTTAGAGCTTTGAGAGCACCAAAATATCTCTCTGTGCG 1860
Db 2061 GCATCGATCTGAGGACCTGGCTTTAGAGCTTTGAGAGCACCAAAATATCTCTCTGTGCG 2120
QY 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920
Db 2121 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 2180
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGCGGAACCG 1980
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGCGGAACCG 2240
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2279

RESULT 5
AX137743
LOCUS
DEFINITION
Sequence 8 from Patent EPI077260.
ACCESSION
AX137743
VERSION
AX137743.1 GI:14273917
KEYWORDS
Mice minute virus
SOURCE
ORGANISM
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
1
Nuesch,J. and Rommelaere,J.
Parvovirus ns1 variants
Patent: EP 1077260-A 8 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
source
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1..2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39993.1"
/db_xref="GI:14273918"
/translation="MAGNAYSDEVLGATNWLKEKSNQVFSFVKFNENVQLNGKDIGW
NSYKELQDELKSLQGAETTWQSEDMEWETTVDEMTKQVFI FDSLKKCLPEVL
NTKNI PPGDVNWFVQHEWKGDOGHCHVLIGKDRSQAGKWRRLNLYVWSRWLVA
CNVQLTPABRIKLREIAEDNEWTLTYKHKOTKDYTKCVLFGNMIAYFLTKKIS
TSPRGGTFLSSDSKWNFLKEGERHLVSKLTDMDRPEVETVITTAQETKRGRI
QTKKEVSIKTKLKVHRTVSPEDMMWOPDSYIEMMAQPOGENLLKNTLEICTLTL
ARTKFDLILEKAETSKLTNFSPLDTRACRIPAFHGMNVYKVCHAI CCLVARQGKR
NTVLFGPASTGSI IAOIAOAVGNVGVNANVNFDFNDCTNKNLI WVEAGNFGQ
OVNQPKAICSGOTIRIDOKGSKJOIEPTVIMTNENITVVRIGCEERPEHTOPRD
RLNLIHLTLFGDFGLVDKNWPMICAMLVKNGYQSTWASVCXKWKVYPDHSNWA
PKVPTINLIGSARSFPTPKSLPNTALPTPLASDLSDLALEPWSPTENTPVAGTAE
TONTGAGSKACODGQLSPTWSEIEBDRACFGAEPLKKDFSEPLNLD"

ORIGIN
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
QY 61 AGTAAACGAGGAAGTGTCTCATTTGTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
```

QY	1201	GCAGCAGCGCAATCTATTATTGACAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
DB	1201	GCAGCAGCGCAATCTATTATTGACAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
QY	1261	TGCTATAATGACGCAATGTAATCTTTCCATTTAATGACTGTACCAACAGAACTTGATT	1320
DB	1261	TGCTATAATGACGCAATGTAATCTTTCCATTTAATGACTGTACCAACAGAACTTGATT	1320
QY	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT	1380
DB	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT	1380
QY	1381	GGTCAAACTATTGCTATGATCAAAAGGAAAGGAGGCAAGAAAGAGTGAACCAACCA	1440
DB	1381	GGTCAAACTATTGCTATGATCAAAAGGAAAGGAGGCAAGAAAGAGTGAACCAACCA	1440
QY	1441	GTCAATCATGACCAATGAGACATTTACAGTGTGAGATAGCTGCGAAGAAAGACCA	1500
DB	1441	GTCAATCATGACCAATGAGACATTTACAGTGTGAGATAGCTGCGAAGAAAGACCA	1500
QY	1501	GAACACACTCAACCAATCAGACAGAGATGCTTAAACATTTCACTAAACATACCTTGCT	1560
DB	1501	GAACACACTCAACCAATCAGACAGAGATGCTTAAACATTTCACTAAACATACCTTGCT	1560
QY	1561	GGTGACTTTGGTTGTTGACAAAATGAATGGCCCAATGATTTGCTGGTTGGTAAAG	1620
DB	1561	GGTGACTTTGGTTGTTGACAAAATGAATGGCCCAATGATTTGCTGGTTGGTAAAG	1620
QY	1621	AATGGTTTACCAATCTACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
DB	1621	AATGGTTTACCAATCTACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
QY	1681	TCAGAAAACCTGGCGGAGCCAAAGGTCGCAACTCTTAAATTTACTAGTTCGGCAGCG	1740
DB	1681	TCAGAAAACCTGGCGGAGCCAAAGGTCGCAACTCTTAAATTTACTAGTTCGGCAGCG	1740
QY	1741	TCACCAATTCAGACACCGAAAGTACGCTCTCAGCGAGAACTATGCACTAACTCCACTT	1800
DB	1741	TCACCAATTCAGACACCGAAAGTACGCTCTCAGCGAGAACTATGCACTAACTCCACTT	1800
QY	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATCTCCTGTTGG	1860
DB	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATCTCCTGTTGG	1860
QY	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGTCAT	1920
DB	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGTCAT	1920
QY	1921	CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAACCG	1980
DB	1921	CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAACCG	1980
QY	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019
DB	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019

RESULT 6  
AX137747  
LOCUS AX137747 2019 bp DNA linear PAT 30-MAY-2001  
DEFINITION Sequence 12 from Patent EP1077260.  
ACCESSION AX137747  
VERSION AX137747.1 GI:14273921  
KEYWORDS Mice minute virus  
SOURCE Mice minute virus  
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
REFERENCE 1  
Nuesch, J. and Rommelaere, J.  
Parvovirus ns1 variants  
Patent: EP 1077260-A 12 21-FEB-2001;  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

(DE)

FEATURES	Location/Qualifiers
source	1. .2019
	/organism="Mice minute virus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:10794"
	/note="Parvovirus NS1 variant"
CDS	1. .2019
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC39995.1"
	/db_xref="GI:14273922"
	/translation="MAGNAYSDVELGATNWLKEKSNQEVFSFVKENVOLNGKDIGW
	NSYKLEQDELKSLQRGAEETWDQSEDMETVDEMTKKQVFIKSLVKKLFEVL
	NTKIPFGDVNMFVQHEWKGQDQGHVLIIGKDFSOAQGKRRQLNVWWSRLVTA
	CNVQTPAERIILREIAEDNFWLLTYKHQTKDYTKCVLFENMIAYFLTKKIS
	TSPRDGGVFLSSDKTFLKEGERHLVSKLYTDDMRPEVETVTTTAQETKRGRI
	OTKKEYSIKTKLKELVHRTVPEDMMWQPDYSIEMMAQPGENLKNLTLEICTLTL
	ARTKPAFDLLEKAEKSLTNFSLPDRTRIFAFHGMVYKVCALCCVLNRQGGKR
	NAVLFPGASTKSIITQAALQAGVNGVCNNAANVPFNDCTNKNLIWVEAGNFGQ
	QVNQFAICSGQIRIDQKGSQKIEPTVPVNTNENITTVRIGCEERPEHTOPIRD
	RMLNIHLTHLPDGLVDKNWPMICAWLVKNGYQSTWASCKAKGVKPDWSENNAE
	PKVPTPINLGSARSPTPKSTPLSQNTVALTPLASDLLELALPEMSTPNTPVAGTAE
	TQNTGEAGSKACQDGLSPTWSEIEEDLRACFGAEPLKKDFSEPLND"

Query Match 99.8%; Score 2015.8; DB 6; Length 2019;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
DB	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
QY	61	AGTAACACGAGAGTGTCTCATTTGTTTAAATAATGAAATGTTCACACTGAATGGA	120
DB	61	AGTAACACGAGAGTGTCTCATTTGTTTAAATAATGAAATGTTCACACTGAATGGA	120
QY	121	GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTA	180
DB	121	GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTA	180
QY	181	CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGG	240
DB	181	CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGG	240
QY	241	GAAATGACCAAAAGCAAGTATTCTTTGTTTAAATAATGAAATGTTCATTTTAA	300
DB	241	GAAATGACCAAAAGCAAGTATTCTTTGTTTAAATAATGAAATGTTCATTTTAA	300
QY	301	GTGCTTTAACAAGAAATATATTCCTGTTGATGTTAAATTTGTTGCAACATGAATG	360
DB	301	GTGCTTTAACAAGAAATATATTCCTGTTGATGTTAAATTTGTTGCAACATGAATG	360
QY	361	GGAAAGACCAAGCTGGCACTGCCATGTAATTTGGAGGAAAGGACTTTAGTCAAG	420
DB	361	GGAAAGACCAAGCTGGCACTGCCATGTAATTTGGAGGAAAGGACTTTAGTCAAG	420
QY	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGGTAAACAG	480
DB	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGGTAAACAG	480
QY	481	TGTAATGTGCAACTAAACACAGCTGAAAGAAATTAATAAGAAATAGCAGAGACA	540
DB	481	TGTAATGTGCAACTAAACACAGCTGAAAGAAATTAATAAGAAATAGCAGAGACA	540
QY	541	GAGTGGGTTACTTACTTACTTATAAGCATAAGCAAAACCAAAAGACTTACCAAG	600
DB	541	GAGTGGGTTACTTACTTACTTATAAGCATAAGCAAAACCAAAAGACTTACCAAG	600
QY	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAG	660
DB	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAG	660



```
QY 661 CCACCAAGACGAGGAGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Db 661 CCACCAAGACGAGGAGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
QY 721 AAAGAAGCGAGCGCCATCTAGTAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
Db 721 AAAGAAGCGAGCGCCATCTAGTAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAACTAAAGCGCGCAGAACTCAAACTAAAAA 840
Db 781 GTTGAACACACAGTAACCACTGCGCAGGAACTAAAGCGCGCAGAACTCAAACTAAAAA 840
QY 841 GAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTCATAAAAAGAGTAACCTCACCAGAG 900
Db 841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTCATAAAAAGAGTAACCTCACCAGAG 900
QY 901 GACTGGATGATGACGACGACAGTACATTGAATGATGCTCAACAGGTGGAGAA 960
Db 901 GACTGGATGATGACGACGACAGTACATTGAATGATGCTCAACAGGTGGAGAA 960
QY 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
QY 1081 ACAAGAACCTGCAAAATTTTGGCTTTTTCATGGCTGGAACATATGTTAAAGTTTCCCATGCT 1140
Db 1081 ACAAGAACCTGCAAAATTTTGGCTTTTTCATGGCTGGAACATATGTTAAAGTTTCCCATGCT 1140
QY 1141 ATTGCTGTGTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTTATTTCATGACCA 1200
Db 1141 ATTGCTGTGTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTTATTTCATGACCA 1200
QY 1201 GCCAGCAGCGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
Db 1201 GCCAGCAGCGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
QY 1261 TGCTATAATGACGCAATGTAACCTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGCTATAATGACGCAATGTAACCTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
QY 1321 TGGGTAGAAGAGCTGTAACCTTTGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Db 1321 TGGGTAGAAGAGCTGTAACCTTTGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
QY 1381 GGTCAAACTATTTCGCAATGATCAAAAGGAAAAAGGAGCAAAACAGATTGAACCAACCA 1440
Db 1381 GGTCAAACTATTTCGCAATGATCAAAAGGAAAAAGGAGCAAAACAGATTGAACCAACCA 1440
QY 1441 GTCATATGACCAAAATGAGAACATTACAGTGGTCAGATAGCTGCGGAGAAAGACCA 1500
Db 1441 GTCATATGACCAAAATGAGAACATTACAGTGGTCAGATAGCTGCGGAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTATCATCTAACACATCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTATCATCTAACACATCTTGCT 1560
QY 1561 GGTGACTTTGGTTGTTGACAAAAATGAATGGCCCATGATTGCTTTGGTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTTGTTGACAAAAATGAATGGCCCATGATTGCTTTGGTTGGTAAAG 1620
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGCTAAATGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGCTAAATGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAAACTGGCGGAGCAAAAGGTGCCAATCTCTATAAATTTACTAGTTCGGCAGCG 1740
Db 1681 TCAGAAAACTGGCGGAGCAAAAGGTGCCAATCTCTATAAATTTACTAGTTCGGCAGCG 1740
```

```
1741 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCGACGAATATGCACTAACTCCACTT 1800
1741 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCGACGAATATGCACTAACTCCACTT 1800
1801 GCATCGGATCTCGAGACCTGGCTTTTAGAGCCTTTGAGACACACCAAAATACTCTGTTGG 1860
1801 GCATCGGATCTCGAGACCTGGCTTTTAGAGCCTTTGAGACACACCAAAATACTCTGTTGG 1860
1861 GGCACCTGCAGAAACCCAGAACACTGGGGAGCTGGTTTCAAAGCCTGCCAAAGATGTCAA 1920
1861 GGCACCTGCAGAAACCCAGAACACTGGGGAGCTGGTTTCAAAGCCTGCCAAAGATGTCAA 1920
1921 CTGAGCCCCAATTTGGTTCAGAGATCGAGAGGATTTTGAGAGCGTGTCTCGGTGCGGAACCG 1980
1921 CTGAGCCCCAATTTGGTTCAGAGATCGAGAGGATTTTGAGAGCGTGTCTCGGTGCGGAACCG 1980
1981 TTGAAGAAAGACTTCAGCGAGCGCTGGAACCTTGGACTAA 2019
1981 TTGAAGAAAGACTTCAGCGAGCGCTGGAACCTTGGACTAA 2019

RESULT 7
AX137751
LOCUS AX137751 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 16 from Patent EP1077260.
ACCESSION AX137751
VERSION AX137751.1 GI:14273925
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Nuesch, J. and Rommelaere, J.
Parvovirus ns1 variants
Patent: EP 1077260-A 16 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)
FEATURES
source
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39997.1"
/db_xref="GI:14273926"
/translation="MAGNAYSDEVILGATNWLKESNQEVFVFKNENVQLNGKDIGW
NSYKLEQDELKSLQGAETTWDSQSEDMENETTVDEMTKKQVFI FDSLVRKCLFEVL
NTKNI FPGDVNMFQHEWKGDKDQCHVLIGKDFSOAGKWRQLNLYWRLVTA
CNVQLTPAERI KLREIAEDNEWVLLTVKHKOTKDYTKVLFGNMIA YVFLTKKIS
TSPPDGGYFLSSDSGKTNELKEGERHLVSKLYTDDMRPETVETVTTAOSTKGRRI
QTKESVSKITLKLHLKRVSPEDMMQPSYIEMMQPGENLLKNTLSICITLTL
ARTKTAFLDILEKETS KLTNPSLPDTRTCRIFAFHGMNVKVCCHAICVLARQGGKR
NTVLFHPSASTKSI IAQIAQAVNGVCYNAANVPFNDCTNKNLIWVBAGNFGQ
QMNQKAIQIDQKGSQIEPTVIMTNNENITVVRIGCEERPEHTQPTRD
RMLNTHLTHTLPGDFGLVDKNEWPMI CAWLKNGYQSTWASCAWKGVDPDSENWAE
PKVPTPINLIGARSPTTPKSTPLSONVALTPLASDLEDLAEPLWSTPNTPVAGTAE
TONTGEGSKACQDQQLSPTWSEIBEDLRACFGAEPLKKDFSEPLNLD"
ORIGIN
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATGCTTACTCTCATGAACTTTTGGGAGCAACCAACTGGTTAAGGAAAA 60
Db 1 ATGGCTGGAAATGCTTACTCTCATGAACTTTTGGGAGCAACCAACTGGTTAAGGAAAA 60
QY 61 AGTAACCGAGAGTCTTCTCATTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 61 AGTAACCGAGAGTCTTCTCATTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
```

121 GATATCGGATGGAATAGTTTCAAAAAGAGCTGCGAGAGACGAGCTGAAATCTTTACAA 180  
121 GATATCGGATGGAATAGTTTCAAAAAGAGCTGCGAGAGACGAGCTGAAATCTTTACAA 180  
181 CGAGGAGCGAAACTACTTCGGACCAAGAGGAGACATGGAATGGGAAACACACAGTGGAT 240  
181 CGAGGAGCGAAACTACTTCGGACCAAGAGGAGACATGGAATGGGAAACACACAGTGGAT 240  
241 GAAATGACCAAAAAGCAAGATTCATTTTGTGTTTAAATAATGTTTAAATTTGAA 300  
241 GAAATGACCAAAAAGCAAGATTCATTTTGTGTTTAAATAATGTTTAAATTTGAA 300  
301 GTGCTTTAACACAAAGAATATATTTCTGGTGATGTTAAATGGTTTGTGGAACATGAATGG 360  
301 GTGCTTTAACACAAAGAATATATTTCTGGTGATGTTAAATGGTTTGTGGAACATGAATGG 360  
361 GGAAGAGACCAAGCTGGCACTGCATGTACTTAATTTGGAGAAAGGACTTTTAGTCAAGCT 420  
361 GGAAGAGACCAAGCTGGCACTGCATGTACTTAATTTGGAGAAAGGACTTTTAGTCAAGCT 420  
421 CAAGGGAAATGTTGGAGAAAGCAACTAAATGTTTTACTTGGAGCAGATGTTGGTAACAGCC 480  
421 CAAGGGAAATGTTGGAGAAAGCAACTAAATGTTTTACTTGGAGCAGATGTTGGTAACAGCC 480  
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 540  
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 540  
541 GAGTGGGTTACTTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600  
541 GAGTGGGTTACTTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600  
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAGAAATAAGCACTAGT 660  
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAGAAATAAGCACTAGT 660  
661 CCAACCAAGAGACGAGGCTATTTTCTTAGCAGTGAATCTCTGGCTGGAAAACTAACTTTTAA 720  
661 CCAACCAAGAGACGAGGCTATTTTCTTAGCAGTGAATCTCTGGCTGGAAAACTAACTTTTAA 720  
721 AAGAGAGCGAGCGCACTAGTGGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
721 AAGAGAGCGAGCGCACTAGTGGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
781 GTTGAACACCAAGTAACTGCGCAGAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840  
781 GTTGAACACCAAGTAACTGCGCAGAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840  
841 GAAATGCTTATTTAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900  
841 GAAATGCTTATTTAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900  
901 GACTGGATGATGCGACGACAGTATTAATGGAATGCTTCAACAGCTGGAGAA 960  
901 GACTGGATGATGCGACGACAGTATTAATGGAATGCTTCAACAGCTGGAGAA 960  
961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAACCAAAACAGCA 1020  
961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAACCAAAACAGCA 1020  
1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGAC 1080  
1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGAC 1080  
1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACATGTTAAAGTTTGCATGCT 1140  
1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACATGTTAAAGTTTGCATGCT 1140  
1141 ATTTCGTGCTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCATGGACCA 1200  
1141 ATTTCGTGCTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCATGGACCA 1200

1201 GCAGACAGGCAAAATCTATTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGGT 1260  
1201 GCAGACAGGCAAAATCTATTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGGT 1260  
1261 TGTATAATGCGCCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACCAAGAACTTGAT 1320  
1261 TGTATAATGCGCCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACCAAGAACTTGAT 1320  
1321 TGGGTAGAAAGAGCTGGTAACTTTTGACACAGCAAGTAAACAGTTTAAAGCATTTGCTCT 1380  
1321 TGGGTAGAAAGAGCTGGTAACTTTTGACACAGCAAGTAAACAGTTTAAAGCATTTGCTCT 1380  
1381 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGGAGCAAGATTTGAACCAACACCA 1440  
1381 GGTCAAGCTATTTCGCATTCATCAAAAGGAAAGGAGGAGCAAGATTTGAACCAACACCA 1440  
1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500  
1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500  
1501 GAACACACTCAACCAATCGAGACAGAACTGCTTAAACATTCATCTAACACATAGCTTGCCT 1560  
1501 GAACACACTCAACCAATCGAGACAGAACTGCTTAAACATTCATCTAACACATAGCTTGCCT 1560  
1561 GGTGACTTTGGTTGGTTGTCACAAAATGATGGCCCATGATTTGCTGCTTGGTTAAAG 1620  
1561 GGTGACTTTGGTTGGTTGTCACAAAATGATGGCCCATGATTTGCTGCTTGGTTAAAG 1620  
1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1680  
1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1680  
1681 TCAGAAAACTGGCGGAGCCAAAGTGCCTTCTTAAATTTACTAGTTCGGCAGCG 1740  
1681 TCAGAAAACTGGCGGAGCCAAAGTGCCTTCTTAAATTTACTAGTTCGGCAGCG 1740  
1741 TCACCAATTCACGACACGAAAGTAGTGCCTCTCAGCCAGAACTATGCACTTAATCCACTT 1800  
1741 TCACCAATTCACGACACGAAAGTAGTGCCTCTCAGCCAGAACTATGCACTTAATCCACTT 1800  
1801 GCATCGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG 1860  
1801 GCATCGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG 1860  
1861 GGCACTGCGAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
1861 GGCACTGCGAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG 1980  
1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG 1980  
1981 TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
1981 TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

## RESULT 8

## MVMICG

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

5085 bp ss-DNA linear VRL 05-JUL-1995  
Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,  
complete genome.

M12032

M12032.1 GI:332289

alternative splicing; capsid protein; complete genome;

nonstructural protein.

Mice minute virus

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

1 (bases 1 to 5085)

Atell,C.R., Gardiner,E.M. and Tattersall,P.

DNA sequence of the lymphotropic variant of minute virus of mice,

MVM(i), and comparison with the DNA sequence of the fibrotropic



ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948; Conservative	0; Mismatches 71; Indels 0; Gaps 0;	
QY	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA	60
DB		
QY	262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAAACCAACCACTGGTTAAAGGAAAA	321
DB		
QY	61 AGTAACCAAGAGTGTCTCAATTTGTTTAAAAATGAAATGTTCAACTGAAATGGAATA	120
DB		
QY	322 AGTAACCAAGAGTGTCTCAATTTGTTTAAAAATGAAATGTTCAACTGAAATGGAATA	381
DB		
QY	121 GATATCGATGGAATAGTTTACAAAAGAGCTGAGAGGACGAGCTGAAATCTTTACAA	180
DB		
QY	382 GATATCGATGGAATAGTTTACAAAAGAGCTGAGAGGACGAGCTGAAATCTTTACAA	441
DB		
QY	181 CGAGGAGCGAAATCTCTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	240
DB		
QY	442 CGAGGAGCGAAATCTCTGGGACCAAGCGAGGACATGGAATGGGAAATCTAGTGGAT	501
DB		
QY	241 GAAATGACCAAAAAGCAAGTATTCATTTTGAATCTTTGTTTAAAAATGTTTATTTGAA	300
DB		
QY	502 GAAATGACCAAAAAGCAAGTATTCATTTTGAATCTTTGTTTAAAAATGTTTATTTGAA	561
DB		
QY	301 GTCTTTAACAACAAGATATTTCTCTGGTGTATTTAAATTTGTTTGTGCAACATGAATGG	360
DB		
QY	562 GTCTTTAGCAAAAAATATAGCTCTCTGCTGATGTTTACTTTGTTGTGACGATGAATGG	621
DB		
QY	361 GGAAGAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTAGTCAAGCT	420
DB		
QY	622 GGAAGAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTAGTCAAGCT	681
DB		
QY	421 CAAGGAAATGTTGGAGAGCAAGCAATTAATTTTACTTGGAGCAGATGTTGGTAAACAGCC	480
DB		
QY	682 CAAGGAAATGTTGGAGAGCAAGCAATTAATTTTACTTGGAGCAGATGTTGGTAAACAGCC	741
DB		
QY	481 TGTAAATGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT	540
DB		
QY	742 TGTAAATGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT	801
DB		
QY	541 GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAACCAAGGACTATACCAAGTGT	600
DB		
QY	802 GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAACCAAGGACTATACCAAGTGT	861
DB		
QY	601 GTTCTTTTGGAAACATGATGTTCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT	660
DB		
QY	862 GTTCTTTTGGAAACATGATGTTCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT	921
DB		
QY	661 CCACCAAGAGACGGAGCTATTTTCTTACGAGTACTCTGGCTGGAAAACTAACTTTTAA	720
DB		
QY	922 CCACCAAGAGACGGAGCTATTTTCTTACGAGTACTCTGGCTGGAAAACTAACTTTTAA	981
DB		
QY	721 AAGAGAGGAGGAGCGCATCTAGTGAGCAACTATATACATGATGACATGCGGCGAGAAACG	780
DB		
QY	982 AAGAGAGGAGGAGCGCATCTAGTGAGCAACTATATACATGATGACATGCGGCGAGAAACG	1041
DB		
QY	781 GTTGAACCACTAATCACTGCGAGGAACTAAGCGGCGAGAAATTAAGCACTAAATAA	840
DB		
QY	1042 GTTGAACCACTAATCACTGCGAGGAACTAAGCGGCGAGAAATTAAGCACTAAATAA	1101
DB		
QY	841 GAAGTGTCTATTAAAACTACATTTAAAGAGCTGGTGATATAAAGAGTAACCTCACAGAG	900
DB		
QY	1102 GAGTGTCTATTAAAACTACATTTAAAGAGCTGGTGATATAAAGAGTAACCTCACAGAG	1161
DB		
QY	901 GACTGATGATGATGAGCAGACAGTTACATTTGAAATGATGCTCAACAGAGTGGAGAA	960
DB		
QY	1162 GACTGATGATGATGAGCAGACAGTTACATTTGAAATGATGCTCAACAGAGTGGAGAA	1221
DB		
QY	961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGCAACCAACAGCA	1020
DB		
QY	1222 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGCAACCAACAGCA	1281
DB		

RESULT 9

MOU34253

LOCUS

MOU34253

4764 bp

DNA

linear

VRL 21-AUG-1996









OTKKEVSIKTTLKELVHKRVTSPEDEMMOPDSYIEWMAQPGSENLLKNTLEICTLTL  
ARTKFDLLEKAEKTSKLTNPSLPTDRCKIKYFAPHGWNITKVKCAICCVLNKQGGKR  
NTVLFHPSGTGSIITQAIAQAVNGVNCYNAANVFPFNCTNKNLIWVEAGNFGQ  
QNOFAIKCSGTIRIDQKGGKQIEPTVIMTNTNENITVVKIGKEERPEHTQPIRD  
RMLNHLHTLPDGFGLVDNKPWMI CAWLKNGYOS TWAS YCAKRGKVPDPTENWAE  
PKVPTPINSIGSARSPTTPKSPLSQNYALTPLASDLEDLALPEWSTNTVPVGTAE  
TQNTGEAGSKACODGQLSPWBSIEBDRALCFAGPEPLKXDFSEPLND"  
join(2290..2320,2403..4561)  
/gene="Vp1"  
CDS  
join(2290..2320,2403..4561)  
/gene="Vp1"  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAA61406.1"  
/db\_xref="GI:525327"

AYDKYIKSGKNPVLVYFSADOREIDOTKADWGGKVGHVFFRTKFAFAPRLASSEP  
GTSGVSIAGRTKPPAHIFINOARAKKRSLAAQOORTOTMSDAGOPDGSVQSA  
RVERAADGPGSGSGSGGGVGVSGVSDNTHYRFLSDGWVEITAYSTRVHLMNP  
KSENYCRVRVHNTNRTAGNMAKDDAHOIWPWSLIDSNWGVFPQSDWFOFINP  
MSHVNLSLDOELFNVIKTVTEONTGAEAVKIYNNDLTASMMVALDSNNILPYTPT  
DNQETLGFYPMKFTSPYRYFNCDSLSVTYDQTSIVDTMANASGLSOPFFIE  
NTORIQLLRTGDFATGTYFETPEPIKLSHTWOSNQLGPOPOITDLPADNENATLV  
TRGDRSGITQISGNSNVTBATRVRPAQVGFQPHDNFETSRAGPVPVVPVNAVTOGN  
EHDANGSLRYTDKQGDWGSNNKERTWDALSYDSGRWADRCFINATPPTSPPAL  
NNLTNDSDPIGNKTAHYQNVFNSYGLPTAFPHAPPIYPOGQIWDKELDLHKKPLRHA  
QAPFVCKNNAPOGLLVRLAPNLTDQYDPSNSTLSRYTYGTFPFWKGKLTAKAKLRPNA  
TWNPVYQVSAQYQNEVENS IHKWLPTATGNMQSIPLLSRPVARNTY"

## ORIGIN

Query Match 94.1%; Score 1899; DB 14; Length 5144;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 1944; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60  
DB 265 ATGGCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 324  
QY 61 AGTAACCAAGAGTGTCTCATTTGTTTTTAAAAATGTAATGTCAACTGAAATGGAAAA 120  
DB 325 AGTAACCAAGAGTGTCTCATTTGTTTTTAAAACTGAGGATGTTCAACTAAATGGAAAA 384  
QY 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAGCTGAAATCTTTACAA 180  
DB 385 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAGCTGAAATCTTTACAA 444  
QY 181 CGAGGAGCGAACTACTTCGGACCAAGAGGAGCATGGAATGGGAACCAAGTGGAT 240  
DB 445 CGAGGAGCGAACTACTTCGGACCAAGAGGAGCATGGAATGGGAATCTACAGTGGAT 504  
QY 241 GAAATGACCAAAAAGCAAGTATTCAATTTTGTATTTTGTGTTAAAAAATGTTTATTGAA 300  
DB 505 GAAATGACCAAAAAGCAAGTATTCAATTTTGTATTTTGTGTTAAAAAATGTTTATTGAA 564  
QY 301 GTGCTTAAACAAAAGAAATATTTCTCGTGTGATGTTAAATTTGTTGTGCAACATGAATGG 360  
DB 565 GTGCTTAGCACAAAGAAATATAGTCTCTGCTGATGTTACTTTGTTGTGCAAGCATGAATGG 624  
QY 361 GGAAGAAGCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGAGCTTTTGTCAAGCT 420  
DB 625 GGAAGAAGCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGAGCTTTTGTCAAGCT 684  
QY 421 CAAGGGAATGTTGGAGGAAGCACTAAATGTTTTTACTTGGAGCAGATGTTGTTGTAACAGCC 480  
DB 685 CAAGGGAATGTTGGAGGAAGCAGCTAAATGTTTTTACTTGGAGCAGATGTTGTTGTAACAGCC 744  
QY 481 TGTAAATGTGCAATTAACACAGCTGAAAGAAATTAACATAAGAAATAGCAGAAAGCAAT 540  
DB 745 TGTAAATGTGAGCTTAACACAGCTGAAAGAAATTAACATAAGAAATAGCAGAAAGCAGT 804  
QY 541 GAGTGGGTACTTCTTACTTATAGCATAAGCAACCAAAAGCAATATACCAAGTGT 600  
DB 805 GAGTGGGTACTTCTTACTTATTAACATPAAGCAAAACCAAAAGGAGCTATATAAATGT 864

QY 601 GTTCTTTTTGGAAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAAAATAGCACTAGT 660  
DB 865 GTTCTTTTTTGGAAATATGATTGCTTACTACTTTTTTAAACCAAAAAAATAAGCACCAGT 924  
QY 661 CCACCAAGAGACGGAGGCTATTTTTCTTAGCAGTGACTCTCGCTGGAAAACTAACTTTTTTA 720  
DB 925 CCGCCAAGAGACGGAGGCTATTTTTCTAAGCAGTGACTCTCGCTGGAAAAACTTAATCTTTTA 984  
QY 721 AAAAGAGGGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATCGCGCCAGAAACG 780  
DB 985 AAAGAGGGCGAACGCCATCTAGTGAGCAAACTATACATCTGATGACATCGCGACCAAGAACG 1044  
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAAAACTAAGCGCGGAGAAATTTCAAACTAAAAA 840  
DB 1045 GTTGAACACACAGTAACCACTGCGCAGGAAAACTAAGCGCGGAGAAATTTCAAACTAAAAA 1104  
QY 841 GAAGTTGCTATTAAAACTACACTTAAAGAGCTGCTGCATAAAAAGAGTAACCTCACAGAG 900  
DB 1105 GAGGTTTCTATTAAAAACCACTTAAAGAGCTGCTGCATAAAAAGAGTAACCTCACAGAA 1164  
QY 901 GACTGGATGATGATGCGAGCCAGACAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960  
DB 1165 GACTGGATGATGATGCGAGCAGACAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAA 1224  
QY 961 AACTGCTGAAAAAATACGCTAGAGATTTGTACATTAACCTTAGCCAGAACCAAAACAGCA 1020  
DB 1225 AACTGCTGAAAAAATACGCTAGAGATCTGTACATAACTCTAGCTAGAACCAAAACAGCA 1284  
QY 1021 TTTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTTCACCTGCCTGAC 1080  
DB 1285 TTTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTTCACCTGCCTGAC 1344  
QY 1081 ACAAGAACCTGCAAGAAATTTTGTCTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT 1140  
DB 1345 ACAAGAACCTGCAAGATCTTTGCTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT 1404  
QY 1141 ATTGCTGTGTTTTTAAACAGACAGAAAGGAGGCAAAAGAAATACTGTTTATTTTATGGAACA 1200  
DB 1405 ATTGCTGTGTTTTTAAACAGACAGAAAGGAGGCAAAAGAAATACTGTTTATTTTATGGAACA 1464  
QY 1201 GCCAGACAGGCAAACTATTTATTGACAGCCATAGCAAGCAGTTGCAATGTTGGT 1260  
DB 1465 GCCAGTACAGGCAAACTTCCATTTATGCAAGCCATAGCAGGAGCTTGTATGTTGGT 1524  
QY 1261 TGTATAATGCAAGCAATGTAAACTTTTCAATTTTAACTGTACCAACCAAGAACTTGAT 1320  
DB 1525 TGTATAATGCAAGCAATGTGAATTTTCCATTTCAATGACTGTACCAACCAAGAACTTAAT 1584  
QY 1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAACCAAGTTTAAAGCAATTTGCTCT 1380  
DB 1585 TGGGTAGAGAAAGCTGGTAACTTTTGGACAAAGTAACCAAGTTTAAAGCAATTTGCTCT 1644  
QY 1381 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGGAGCAAGGATTGAACCAACACCA 1440  
DB 1645 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGGAGCAAGGATTGAACCAACACCA 1704  
QY 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGCTCAGAAATAGGCTGCGAAGAAAGACCA 1500  
DB 1705 GTCATCATGACCAAAATGAGAACTTACAGTGGCTCAGAAATAGGCTGCGAAGAAAGACCA 1764  
QY 1501 GAAACACTCAACCAATCAGAGACAGAAATGCTTAACTTCAATCTAACACATACATCTGCCT 1560  
DB 1765 GAAACACTCAACCAATCAGAGACAGAAATGCTTAACTTCAATCTAACACATACATCTGCCT 1824  
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGATGGCCCATGATTTGCTGCTGTTGGTTAAAG 1620  
DB 1825 GGTGACTTTGGTTGGTTGACAAAAATGATGGCCCATGATTTGCTGCTGTTGGTTAAAG 1884  
QY 1621 AATGGTTTACCAATCTACCACTGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTTCTGATTGG 1680  
DB 1885 AATGGTTTACCAATCTACCACTGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTTCTGATTGG 1944

QY	1681	TCAGAAACTGGCGGAGCCAAAGGTGCCAATCTCTATAAATTTACTAGTTCGGACGC	1740
Db	1945	ACGAAAACCTGGCGGAGCCGAAGGTGCCAATCTCTATAAATTTCACTAGTTCGGACGC	2004
QY	1741	TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACATGACACTAACTCCACTT	1800
Db	2005	TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACATGACACTAACTCCACTT	2064
QY	1801	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG	1860
Db	2065	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG	2124
QY	1861	GGCACTCGAGAACCCAGAACACTGCGGGAAGTGGTTCCAAAGCTGCGCAAGATGTCFAA	1920
Db	2125	GGCACTCGAGAACCCAGAACACTGCGGGAAGTGGTTCCAAAGCTGCGCAAGATGTCFAA	2184
QY	1921	CTGAGCCCAACTGGTCAGAGATCGAGGAGATTCAGAGCGTTCGCTGCGGAACCG	1980
Db	2185	CTGAGCCCAACTGGTCAGAGATCGAGGAGATTCAGAGCGTTCGCTGCGGAACCG	2244
QY	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019
Db	2245	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2283
RESULT 12			
MOU34254			
LOCUS			
Mouse parvovirus 1c DNA.			
ACCESSION			
U34254.1			
VERSION			
GI:1464794			
KEYWORDS			
SOURCE			
ORGANISM			
Mouse parvovirus 1c			
Mouse parvovirus 1c			
REFERENCE			
1 (bases 1 to 4764)			
Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.B., Hook, R.R. Jr. and Riley, L.K.			
Molecular characterization of newly recognized rodent parvoviruses			
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)			
JOURNAL			
MEDLINE			
PUBMED			
8609486			
REFERENCE			
2 (bases 1 to 4764)			
Besselsen, D.G.			
Direct Submission			
Submitted (17-AUG-1995) David G. Besselsen, Department of			
Veterinary Pathology, University of Missouri-Columbia, W213			
Veterinary Medicine Building, Columbia, MO 65211, USA			
JOURNAL			
FEATURES			
source			
1. 4764			
/organism="Mouse parvovirus 1c"			
/mol_type="genomic DNA"			
/specific_host="Mus musculus"			
/db_xref="taxon:42842"			
ORIGIN			
Query Match 92.8%; Score 1873; DB 14; Length 4764;			
Best Local Similarity 95.4%; Pred. No. 0;			
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;			
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Db	122	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCACTGGTTAAAGGAAAA	181
QY	61	AGTAAACAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
Db	182	AGTAAACAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	241
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGAGCTGAAATCTTTACAA	180
Db	242	GATATTTGGATGGAATATTACAAAGAGAGCTGCAAGAGAGCAGAGCTAAAAATCTTTACAA	301
QY	181	CGAGGAGCGGAACACTACTTGGGACCAAGAGGAGCATGGAATGGGAAACCAAGTGGAT	240

Db	302	CGAGGAGCGGAACACTACTTGGGACCAAGGAGCATGGAATCTACAGTGGAT	361
QY	241	GAAATCACCACCAAAAGCAAGTATTTCATTTTGGTTGTTTAAAAAATGTTTATTGAA	300
Db	362	GAAATGACCAAAAGCAAGTATTTCATTTTGGTTGTTTAAAAAATGTTTATTGAA	421
QY	301	GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTTAAATTTGGTTTGTGCAACATGAATGG	360
Db	422	GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTTAAATTTGGTTTGTGCAACATGAATGG	481
QY	361	GGAAAGACCAAGCTGGCACTGCGCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420
Db	482	GGAAAGACCAAGCTGGCACTGCGCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT	541
QY	421	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGTTGTTAAACAGCC	480
Db	542	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGTTGTTAAACAGCC	601
QY	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
Db	602	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	661
QY	541	GAGTGGTTTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT	600
Db	662	GAGTGGTTTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT	721
QY	601	GTTCCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAAAATAGCACTAGT	660
Db	722	GTTCCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAAAATAGCACTAGT	781
QY	661	CCACCAAGAGACGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	720
Db	782	CCGCCAAGGACGAGGCTTATTTCTTAGTACTGCTGGCTGGAAAACTAACTTTTAA	841
QY	721	AAAGAGGCGAGCGGCATCTAGTAGCAAACTATACACTGATGACATGCGGCGCAGAAACG	780
Db	842	AAAGAGGCGAGCGGCATCTAGTAGCAAACTATACACTGATGACATGCGGCGCAGAAACG	901
QY	781	GTGTAACACACAGTAACTCCTCGCAGGAAACTAAAGCGCGCAGAAATCAAACTAAAAAA	840
Db	902	GTGTAACACACAGTAACTCCTCGCAGGAAACTAAAGCGCGCAGAAATCAAACTAAAAAA	961
QY	841	GAAATGTGCTTATTAACCTAACCTTAAAGAGCTGGTGCATATAAAGAGTAACTCACAGAG	900
Db	962	GAAATGTGCTTATTAACCTAACCTTAAAGAGCTGGTGCATATAAAGAGTAACTCACAGAG	1021
QY	901	GACTGATGATGATGCGCAGCAGACAGTTACATTTGAAATGATGCTCAACAGCTGGAGAA	960
Db	1022	GACTGATGATGATGCGCAGCAGACAGTTATATTTGAAATGATGCTCAACAGCTGGAGAA	1081
QY	961	AACCTGCTGAAAAATACCGCTAGAGATTTGTACACTTAACCTAGCCAGAACCAAAACAGCA	1020
Db	1082	AACCTGCTGAAAAATACCGCTAGAGATTTGTACACTTAACCTAGCCAGAACCAAAACAGCA	1141
QY	1021	TTTGAATTAATTTTAAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCTGAC	1080
Db	1142	TTTGAATTAATTTTAAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCTGAC	1201
QY	1081	ACAAGAACCTGAGAAATTTTTCCTTTTCATGCTGGAATATGTTAAAGTTTGCATGCT	1140
Db	1202	ACAAGAACCTGAGAAATTTTTCCTTTTCATGCTGGAATATGTTAAAGTTTGCATGCT	1261
QY	1141	ATTTCGTGTTTAAACAGACAGGAGCAAAAGAAATACCTGTTTATTTTCATGGACCA	1200
Db	1262	ATTTCGTGTTTAAACAGACAGGAGCAAAAGAAATACCTGTTTATTTTCATGGACCA	1321
QY	1201	GCCAGCAGGCAAAATCTATTATTCACAAGCCATAGCAGACAGTGGCAATGTTGT	1260
Db	1322	GCCAGTACAGGCAAAATCTATTATTCACAAGCCATAGCAGACAGTGGCAATGTTGT	1381
QY	1261	TGCTATATGACCAATGTAACCTTTCATTTATGCTGTAACCAAGACACTGATTT	1320

Db 1382 TGTATATATGACGACATGTGAATCTTCCATTCAATGACTGTACCAACAGAACTTAATT 1441  
Qy 1321 TGGGTAGAGAGCTGTAACTTTGGACAGCAAGTAAACAGATTTAAAGCCATTGTCTCT 1380  
Db 1442 TGGGTGGAAGAGCTGTAACTTTGGACAGCAAGTAAACAGATTTAAAGCCATTGTCTCT 1501  
Qy 1381 GGTCAAACTATTCGCAATGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACACCA 1440  
Db 1502 GGTCAAACTATTCGCAATGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACACCA 1561  
Qy 1441 GTCATCATGACCAACAAATGAGAACATTACAGTGTGTGAGATAGGCTGCGAAGAAAGACCA 1500  
Db 1562 GTCATCATGACCAACAAATGAGAACATTACAGTGTGTGAGATAGGCTGCGAAGAAAGACCA 1621  
Qy 1501 GAACACACTCAACCAATCAGACAGAGATGCTTAAACATTCATCTAACACATACATCTGCT 1560  
Db 1622 GAACACACTCAACCAATCAGACAGAGATGCTTAAACATTCATCTAACACATACATCTGCT 1681  
Qy 1561 GGTGACTTTGGTGTGAGCAAAATGAATGCGCCATGATTTGTGCTTGGTGGTAAAG 1620  
Db 1682 GGTGACTTTGGTGTGAGCAAAATGAATGCGCCATGATTTGTGCTTGGTGGTAAAG 1741  
Qy 1621 AATGGTTTACCAATCTACATGCGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680  
Db 1742 AATGGTTTACCAATCTACATGCGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1801  
Qy 1681 TCAGAAAATCGGGCGAGCCAAAGGTGCCAACTCTTAAATTTACTAGTTTCGGACGC 1740  
Db 1802 ACGGAAAATCGGGCGAGCCAAAGGTGCCAACTCTTAAATTTACTAGTTTCGGACGC 1861  
Qy 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCAGAACTATGCACTTAACCTCACTT 1800  
Db 1862 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCAGAACTATGCACTTAACCTCACTT 1921  
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCTGTTGCG 1860  
Db 1922 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCTGTTGCG 1981  
Qy 1861 GGCACCTCAGAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTTCAA 1920  
Db 1982 GGCACCTCAGAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTTCAA 2041  
Qy 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGACGCTGCTTCGTTGGGAACCG 1980  
Db 2042 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAAAGCGTGTCTCGTTGGGAACCG 2101  
Qy 1981 TTCAAGAAAGACTTCAGCGAGCCGCTGAATCTGACTAA 2019  
Db 2102 TTGAAGAGACTTCAGCGAGCCGCTGAATCTGACTAA 2140

RESULT 13  
HOU34255  
LOCUS HOU34255 4773 bp DNA linear VRL 21-AUG-1996  
DEFINITION Hamster parvovirus DNA.  
ACCESSION U34255  
VERSION U34255.1 GI:1464792  
KEYWORDS  
SOURCE Hamster parvovirus  
ORGANISM Hamster parvovirus  
REFERENCE 1 (bases 1 to 4773)  
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,  
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.  
Molecular characterization of newly recognized rodent parvoviruses  
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)  
96201434  
PUBMED 8609486  
REFERENCE 2 (bases 1 to 4773)  
Besselsen,D.G.  
Direct Submission  
Submitted (17-AUG-1995) David G. Besselsen, Department of  
Veterinary Pathology, University of Missouri-Columbia, W213

Veterinary Medicine Building, Columbia, MO 65211, USA  
Location/Qualifiers  
1. 4773  
/organism="Hamster parvovirus"  
/mol\_type="genomic DNA"  
/specific\_host="Mesocricetus auratus"  
/db\_xref="taxon:42843"  
ORIGIN  
Query Match 92.2%; Score 1862.2; DB 14; Length 4773;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTTAAAGGAAAA 60  
Db 122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAAACCAACCACTGTTTAAAGGAAA 181  
Qy 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGMAAATGTTCAACTGAATGGAAAA 120  
Db 182 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGMAAATGTTCACTCAATGGAAAA 241  
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGCAGAGGACGAGCTGAAATCTTTACAA 180  
Db 242 GATATCGGATGGAATAGTTTACAAAAAGGAGCTCGAAGAGGAGCTGAAATCTTTACAA 301  
Qy 181 CGAGGAGCGGAACTACTTTCGGACCAAGCGAGAGCATGGAATGGGAAACCACTGCTGAT 240  
Db 302 CGAGGAGCGGAACTACTTTCGGACCAAGCGAGAGCATGGAATGGGAAATCTTCACTGCT 361  
Qy 241 GAAATGACCAAAAGCAAGTATTCTTTTTCCTGCTGATGTTAAATTTGTTGTTTAAATGTTT 300  
Db 362 GAAATGACCAAAAGCAAGTATTCTTTTTCCTGCTGATGTTAAATTTGTTGTTTAAATGTTT 421  
Qy 301 GTGCTTAAACACAAAGAAATATATTTTCCTGCTGATGTTAAATTTGTTGTTGTTGTTGTTG 360  
Db 422 GTGCTGATGTTAAACAAAGAAATATATTTTCCTGCTGATGTTAAATTTGTTGTTGTTGTTG 481  
Qy 361 GGAAGAGACCAAGCTTGGACATGCCATGTTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTG 420  
Db 482 GGAAGAGACCAAGCTTGGACATGCCATGTTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTG 541  
Qy 421 CAAAGGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 480  
Db 542 CAAAGGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 601  
Qy 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAAATTAAGAGAAATAGCAGAGCAAT 540  
Db 602 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAAATTAAGAGAAATAGCAGAGCAAT 661  
Qy 541 GAGTGGTTTACTCTTACTTATAAGCATTAAGCAAAACCAAAAGAAAGTATATCAAGTGT 600  
Db 662 GAATGGTCACTTTGCTTACTTATAAGCATTAAGCAAAACCAAAAGAAAGTATATCAAGTGT 721  
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAATTAAGCACTAGT 660  
Db 722 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAATTAAGCACTAGT 781  
Qy 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGCTGTTGTTGTTGTTGTTGTTGTTGTT 720  
Db 782 CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGCTGTTGTTGTTGTTGTTGTTGTTGTT 841  
Qy 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATATACATGATGATGCGGCGAGAAACG 780  
Db 842 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATATACATGATGATGCGGCGAGAAACG 901  
Qy 781 GTTGAACCAACAGTAACCACTGCGGAGGAACTTAAGCGGCGAGAACTTCAACTTAAAAA 840  
Db 902 GTTGAACCAACAGTAACCACTGCGGAGGAACTTAAGCGGCGAGAACTTCAACTTAAAAA 961  
Qy 841 GAAGTTGCTATTAAAACTACACTTAAAGAGCTGCTGCATAAAAAGAGTAACTCACCAGAG 900  
Db 962 GAGGTCTATTAAAAACCACTTAAAGAGCTGCTGCATAAAAAGAGTAACTCACCAGAA 1021

Qy	901	GACTGGATGATGATGACGCCAGACAGTTTACATTGAAATGATGGCTCAACAGAGTGGAGAA	960
Db	1022	GACTGGATGATGATGACGCCAGACAGTTTACATTGAAATGATGGCTCAACAGAGTGGAGAA	1081
Qy	961	AACCTGCTGAAAATACGCTAGAGATTGTACACTTAAGCTAGCCAGAAACCAAAACAGCA	1020
Db	1082	AACCTGCTGAAAATACGCTAGAGATTGTACACTTAAGCTAGCCAGAAACCAAAACAGCA	1141
Qy	1021	TTTGACTTAATTTTATAGAAAAAGCTGAAACACAGCAAACTAAACCAACTTTTCACTGCGCTGAC	1080
Db	1142	TTTGACTTAATTTTATAGAAAAAGCTGAAACACAGCAAACTAAACCAACTTTTCACTGCGCGAC	1201
Qy	1081	ACAAGAACCTGCAGAAATTTTGTCTTTTATGCGTGGAACTATGTTAAAGTTTGCATGCT	1140
Db	1202	ACAAGAACCTGCAGAAATTTTGTCTTTTATGCGTGGAACTATGTTAAAGTTTGCATGCT	1261
Qy	1141	ATTTGCTGTGTTTTAAACAGACAGGAGGCGAAAGAAATACTGTTTTTATTTTCATGGACCA	1200
Db	1262	ATTTGCTGTGTTTTAAACAGACAGGAGGCGAAAGAAATACTGTTTTTATTTTCATGGACCA	1321
Qy	1201	GCCAGCACAGGCAAAATCTATTATTGCAACAGCCATAGCACAAAGCAGTTGGCAATGTTGGT	1260
Db	1322	GCCAGTACAGGCAAAATCCATTATTGCAACAGCCATAGCACAAAGCAGTTGGTAAATGTTGGT	1381
Qy	1261	TGCTATAATGCAGCAAAATGTAAACTTTCCATTTTAATGACTGTACCAACAAAGAACTTGATT	1320
Db	1382	TGCTATAATGCAGCAAAATGTAAACTTTCCATTTTAATGACTGTACCAACAAAGAACTTGATT	1441
Qy	1321	TGGGTAGAAAGCTGTTAACTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380
Db	1442	TGGGTAGAAAGCTGTTAACTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1501
Qy	1381	GGTCAAACTATTTCGCATTGTATCAAAAAGGAAAGGACAGCAAAACAGATTGAACCAACACCA	1440
Db	1502	GGTCAAACTATTTCGCATTGTATCAAAAAGGAAAGGACAGCAAAACAGATTGAACCAACACCA	1561
Qy	1441	GTCAATCATGACCAAAATGAGAACATTTACAGTGGTGCAGAAATAGGCTCGGAGAAAGACCA	1500
Db	1562	GTCAATCATGACCAAAATGAGAACATTTACAGTGGTGCAGAAATAGGCTCGGAGAAAGACCA	1621
Qy	1501	GAAACACACTCAACCAATCAGACAGAGAATGCTTTAAACATTCATCTAAACACATACCTTGCT	1560
Db	1622	GAAACACACTCAACCAATCAGACAGAGAATGCTTTAAACATTCATCTTTAAACACATACCTTGCT	1681
Qy	1561	GGTGACTTTTGTTTGGTTTGACAAAATGAAATGGCCCAATGATTTGCTGCTGGTTGGTAAAG	1620
Db	1682	GGTGACTTTTGTTTGGTTTGACAAACATGAAATGGCCCAATGATTTGCTGCTGGTTGGTAAAG	1741
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTTCCTGATTGG	1680
Db	1742	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTTCCTGATTGG	1801
Qy	1681	TCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATTAATTTTACTTAGGTTGGGCAAGC	1740
Db	1802	ACGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATTAATTTTACTTAGGTTGGGCAAGC	1861
Qy	1741	TCACCATTTACAGCACCCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCCACTT	1800
Db	1862	TCACCATTTACAGCACCCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCCACTT	1921
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACCAAAATACTCCTGTTGCG	1860
Db	1922	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACCAAAATACTCCTGTTGCG	1981
Qy	1861	GGCATCTGAGAAAACCCAGAACACATGGGGAGAGCTGGTTCCAAAGCCTGCAAGATGGTCAA	1920
Db	1982	GGCATCTGAGAAAACCCAGAACACATGGGGAGAGCTGGTTCCAAAGCCTGCAAGATGGTCAA	2041
Qy	1921	CTGAGCCCCAATCTGCTCAGAGATCGAGAGGATTTGAGAGCGTCTCTCGGTGCGGAACCG	1980
Db	2042	CTGAGCCCCAATCTGCTCAGAGATCGAGAGGATTTGAGAGCGTCTCTCGGTGCGGAACCG	2101
Qy	1981	TTGAGAAAGAACTTTCAGCGAGCCGCTGAACTTGGACTAA	2019

Db 2102 TTGAAGAGACTTCACGCGCCGCTGAACCTTGCACTAA 2140  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14  
MVU34256 4761 bp DNA linear VRL 21-AUG-1996  
LOCUS  
DEFINITION  
Accession  
Version  
Keywords  
Source  
Organism  
Reference  
Authors  
Title  
Journal  
Medline  
PubMed  
Reference  
Besselsen,D.G.  
Direct Submission  
Submitted (17-AUG-1995) David G. Besselsen, Department of  
Veterinary Pathology, University of Missouri-Columbia, W213  
Veterinary Medicine Building, Columbia, MO 65211, USA  
Location/Qualifiers  
source  
1..4761  
/organism="Mice minute virus"  
/mol\_type="genomic DNA"  
/strain="Cutter"  
/specific\_host="Mesocricetus auratus"  
/db\_xref="taxon:10794"

ORIGIN  
  
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;  
Best Local Similarity 94.9%; Pred. No. 0;  
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAATGCTTACTCTGTATGATGAGTTTCGGGAGCAACCACCTGGTTAAAGGAAAAA 60  
|||  
Db 122 ATGGCTGGAAATGCTTACTCTGTATGAGTTTCGGGAGCAACCACCTGGTTAAAGGAAAAA 181  
|||  
Qy 61 AGTAACCAAGGAGTGTCTCATTTGTTTTTAAAAATGAAAATGTTCAAATGAAATGAAAA 120  
|||  
Db 182 AGTAACCAAGTAGTATTCTCATTTGTTTTTAAAAATGAAGATGTTCAATGAAATGAAAA 241  
|||  
Qy 121 GATATCGGATGGAAATAGTTTACAATAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180  
|||  
Db 242 GATATCGGATGGAAATAGTTTACAATAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 301  
|||  
Qy 181 CGAGGAGCGAAATCTACTTGGGCCAACGAGGACATGGAATGGGAAACCAAGATGAT 240  
|||  
Db 302 CGAGGAGCGAAATCTACTTGGGCCAACGAGGACATGGAATGGGAAATCTTCAATGAT 361  
|||  
Qy 241 GAATGACCAAAAGCAAGTATTTCATTTTGGATTCTTTGGTTAAAAAATGTTTATTTGAA 300  
|||  
Db 362 GAATGACCAAAAGCAAGTATTTCATTTTGGATCTTTTGTATTTAAAGTGTATTGAA 421  
|||  
Qy 301 GTGCTTAAACAAGAATATATTTCCCTGGTGATGTTTAAATGGTTTGTGCAACATGAATGG 360  
|||  
Db 422 GTGCTTAAACAAGAATATATTTCCCTGGTGATGTTTAAATGGTTTGTGCAACATGAATGG 481  
|||  
Qy 361 GGAAGAACCAAGCTGGCACTGCCATGTACTAATTTGGGAAAGGACTTTAGTCAAGCT 420  
|||  
Db 482 GGAAGAACCAAGCTGGCACTGCCATGTACTAATTTGGGAAAGGACTTTAGTCAAGCT 541  
|||  
Qy 421 CAAGGGAATGCTGGGAAAGCAACTAAATGTTTACTTGGGAGCAGATGTTTGGTAAACAGCC 480  
|||  
Db 542 CAAGGGAATGCTGGGAAAGCAACTAAATGTTTACTTGGGAGCAGATGTTTGGTAAACAGCC 601  
|||  
Qy 481 TGTAATGTGCAACTTAACAACAGCTGGAAGAAATTAACATGAAGAAATAGCAGAAACAAT 540  
|||



462	CGAGGAGGGAACTACTCTTGGGACCAAGCGAGACATGGAAATGGGAAACACACAGTGGAT	521
241	GAATGACCAAAAAAGCAAGTATTCATTTTTCGATCTTTTGGTTAAAAAATGTTTATTTGAA	300
522	GAATGACCAAAAAAGCAAGTATTCATTTTTCGATCTTTTGGTTAAAAAATGTTTATTTGAA	581
301	GTGCTTAAACAAGAAATATATTTTCTGGTGATGTTAAATGGTTGGTTGTGCACATGAATGG	360
582	GTGCTTAAACAAGAAATATATTTTCTGGTGATGTTAAATGGTTGGTTGTGCACATGAATGG	641
361	GGAAAAAGCAAGGCTGGCACTGCCATGTACTAAATGGAGGAAAAAGGACTTTTAGTCAAGCT	420
642	GGAAAAAGCAAGGCTGGCACTGCCATGTACTAAATGGAGGAAAAAGGACTTTTAGTCAAGCT	701
421	CAAGGGAAATGCTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	480
702	CAAGGGAAATGCTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	761
481	TGTAATGTGCAACTAAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAAGACAAT	540
762	TGTAATGTGCAACTAAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAAGACAAT	821
541	GAGTGGTTTACTCTACTTTACTTAAGCATAAAGCAAAACCAAAAAAGACTATACCAAGTGT	600
822	GAGTGGTTTACTCTACTTTACTTAAGCATAAAGCAAAACCAAAAAAGACTATACCAAGTGT	881
601	GTTCCTTTTTCGAAACATGATGCTTACTATTTTCTTAACTAAAAAGAAAAAATGAACACTAGT	660
882	GTTCCTTTTTCGAAACATGATGCTTACTATTTTCTTAACTAAAAAGAAAAAATGAACACTAGT	941
661	CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
942	CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	1001
721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATCGCGGCGAGAAACG	780
1002	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTCTGATATCTGATGAGATGAAACCAAGAAACG	1061
781	GTGTGAAACCAAGTAAACCACTGCGCGAGAAACTAAGCGCGGAGAAATCAAACTTAAAAAA	840
1062	GTGAGACCAAGTGACCACTGACAGAGAGCTAAGCGCGGAGAAATCAAACTGATAGAG	1121
841	GAAGTGTCTATTAATACTACATTAAGAGCTGGTGCAATAAGAGGTAACTCAACAGAG	900
1122	GAGGTCTCGATTAATAACCACTCAAAAGAGTTGGTACATAAAAGAGTAACTCAACAGAA	1181
901	GACTTGGATGATGACAGCAGACAGTTACATTCGAAATGATGGCTCAACAGAGTGGAGAA	960
1182	GACTTGGATGATGACAGCAGACAGTTACATTCGAAATGATGGCTCAACAGAGTGGAGAA	1241
961	AACTCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA	1020
1242	AACTCTGCTGAAAAATACACTAGAGATCTGTACACTGACTCTAGCAAGAAACCAAAACAGCC	1301
1021	TTTTCAGCTTAATTTTAGAAGAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCCTGAC	1080
1302	TTTTCAGCTTGAATCTGGAAGAAAGCTGAAACCAAGCAAACTAGCCAACTTTTCCATGGCTAGC	1361
1081	ACAAGAACCTGCAGAAATTTTTCGCTTTTTCATGGCTGGAACTATGTAAAGTTTTCGCATGCT	1140
1362	ACCAGAACCTGTAGAATCTTTTTCGCTGACATGGCTGGAACTATATTAAGCTGCCATGCC	1421
1141	ATTTCGTGTGTTTAAACAGACAGGAGGCAAAAGAAATATCTGTTTTATTTCATGGAACCA	1200
1422	ATCTGTGTGTGTGTAATAGACAAGGAGGCAAAAGGAACACTGTGCTCTTTCAAGGACCA	1481
1201	GCCAGACAGGCAAACTCTATTATTTCACAAGCCCATAGCACAAAGCAGTTGGCAATGTTGCT	1260
1482	GCCAGACAGGCAAACTCTATTATTTCACAAGCCCATAGCACAAAGCAGTTGGTAATGTTGCT	1541
1261	TGCTATAATGCAGCCAAATGTAACCTTTCCATTTTAATGACTGTGTACCAACAAAGAACTTGAT	1320
1542	TGTTACAAATGCTGCAATGTGAATTTTCCATTTTAATGACTGTGTACCAACAAAAATCTGAT	1601

Search completed: January 22, 2005, 21:19:15  
Job time : 8548.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds  
(without alignments)  
11150.532 Million cell updates/sec

Title: us-10-069-056-4  
Perfect score: 2019  
Sequence: 1 atggcgtggaatgcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2019	100.0	2019	5	AAD02799 Parvoviru
2	2017.4	99.9	2019	5	AAD02797 Parvoviru
3	2017.4	99.9	5149	12	ADG39767 Minute vi
4	2015.8	99.8	2019	5	AAD02801 Parvoviru
5	2015.8	99.8	2019	5	AAD02805 Parvoviru
6	2015.8	99.8	2019	5	AAD02803 Parvoviru
7	1040.4	51.5	5049	2	AAT15311 Non-atten
8	1040.4	51.5	5049	2	AAT15312 Attenuate
9	1040.4	51.5	5049	2	AAT88324 Attenuate
10	1038.8	51.5	5049	2	AAT88321 Attenuate
11	1038.8	51.5	5049	2	AAT88320 Canine pa
12	902.6	44.7	3524	1	AAN40252 Sequence
13	307.4	15.2	374	6	ABQ95626 Tumour su
14	285.8	14.2	421	6	ABQ94779 Tumour su
15	284.8	14.1	423	6	ABQ94724 Tumour su
16	284.8	14.1	464	6	ABQ94719 Tumour su
17	284.8	14.1	473	6	ABQ94716 Tumour su
18	284.8	14.1	486	6	ABQ94707 Tumour su
19	284.8	14.1	491	6	ABQ94777 Tumour su
20	284.2	14.1	420	6	ABQ94787 Tumour su
21	283.8	14.1	343	6	ABQ94737 Tumour su

22	283.8	14.1	343	6	ABQ94739	Tumour su
23	283.8	14.1	403	6	ABQ94782	Tumour su
24	283.2	14.0	497	6	ABQ94718	Tumour su
25	280.4	13.9	340	6	ABQ94732	Tumour su
26	278.2	13.8	342	6	ABQ94765	Tumour su
27	273.8	13.6	420	6	ABQ94781	Tumour su
28	269.6	13.4	324	6	ABQ94740	Tumour su
29	263.4	13.0	339	6	ABQ95625	Tumour su
30	261.6	13.0	424	6	ABQ94778	Tumour su
31	163.8	8.1	468	6	ABQ94873	Tumour su
32	162.2	8.0	451	6	ABQ94868	Tumour su
33	162.2	8.0	457	6	ABQ94867	Tumour su
34	162.2	8.0	465	6	ABQ94874	Tumour su
35	162.2	8.0	469	6	ABQ94869	Tumour su
36	162.2	8.0	472	6	ABQ94880	Tumour su
37	162.2	8.0	515	6	ABQ94793	Tumour su
38	162.2	8.0	516	6	ABQ94807	Tumour su
39	162.2	8.0	516	6	ABQ94794	Tumour su
40	162.2	8.0	516	6	ABQ94800	Tumour su
41	162.2	8.0	517	6	ABQ94811	Tumour su
42	162.2	8.0	530	6	ABQ94791	Tumour su
43	162.2	8.0	530	6	ABQ94799	Tumour su
44	161	8.0	516	6	ABQ94796	Tumour su
45	160.6	8.0	530	6	ABQ94805	Tumour su

## ALIGNMENTS

## RESULT 1

AAD02799 standard; DNA; 2019 BP.

AC AAD02799;

DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)

DE Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
XX tumoural disease; gene therapy; mutant; mutein; variant; ds.

OS Parvovirus.  
OS Synthetic.

PH Key Location/Qualifiers  
FT CDS 1..2019

FT mutation /product= "NS1 variant (S283A) protein"  
FT replace(847, A)  
FT /\*tag= b

XX EPI077260-AL.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

XX WPI; 2001-212717/22.

XX P-PSDB; AAY72704.

XX Novel parvovirus non-structure protein variant, useful for treating  
XX tumoral diseases, has a shifted equilibrium between DNA replication and  
XX transcription activities, and cytotoxic activity.

XX Claim 7; Page 11-14; 41pp; English.



XX The present sequence is a DNA encoding parvovirus non-structure 1 protein  
CC (NS1) variant (S283A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoural diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;

Query Match 100.0%; Score 2019; DB 5; Length 2019;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
DB 1 ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGGAATGAAAA 120  
DB 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGGAATGAAAA 120  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAAGCTGTAATCTTTACAA 180  
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAAGCTGTAATCTTTACAA 180  
QY 181 CGAGGAGCGGAATCTTGGGACCAAGCGAGGACATCGAATGGAACCAACAGTGGAT 240  
DB 181 CGAGGAGCGGAATCTTGGGACCAAGCGAGGACATCGAATGGAACCAACAGTGGAT 240  
QY 241 GAAATGACCAAAAAAGCAAGATTCATTTTGTATTTCTTGGTTAAAAATGTTTATTTGAA 300  
DB 241 GAAATGACCAAAAAAGCAAGATTCATTTTGTATTTCTTGGTTAAAAATGTTTATTTGAA 300  
QY 301 GTGCTTTAACAAAAAGAAATATTTCTGTTGATGTTTAAATGTTTGTGCAACATGAATGG 360  
DB 301 GTGCTTTAACAAAAAGAAATATTTCTGTTGATGTTTAAATGTTTGTGCAACATGAATGG 360  
QY 361 GGAAGAGCCAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGCACTTGTCTCAAGCT 420  
DB 361 GGAAGAGCCAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGCACTTGTCTCAAGCT 420  
QY 421 CRAAGGAAATGTTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGGTTGTAACAGCC 480  
DB 421 CRAAGGAAATGTTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGGTTGTAACAGCC 480  
QY 481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATCTAAGAGAAATAGCAGAGCAAT 540  
DB 481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATCTAAGAGAAATAGCAGAGCAAT 540  
QY 541 GAGTGGGTTACTCTACTTACTTATAGCATAAGCAACCAAAAAAGCACTATACCAAGTGT 600  
DB 541 GAGTGGGTTACTCTACTTACTTATAGCATAAGCAACCAAAAAAGCACTATACCAAGTGT 600  
QY 601 GTTCTTTTTGGAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAAAATAGCACTAGT 660  
DB 601 GTTCTTTTTGGAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAAAATAGCACTAGT 660  
QY 661 CCAACAGAGACGGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 720  
DB 661 CCAACAGAGACGGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 720  
QY 721 AAAGAGGGGAGCGGCATCTAGTGAGCAACTATACACTGATGACATGGGCCAGAGAACG 780  
DB 721 AAAGAGGGGAGCGGCATCTAGTGAGCAACTATACACTGATGACATGGGCCAGAGAACG 780  
QY 781 GTTGAACACCACTAATCACTGGCGAGGAACTAAGCGCGGCAAGATTTCAAACTAAAAAA 840  
DB 781 GTTGAACACCACTAATCACTGGCGAGGAACTAAGCGCGGCAAGATTTCAAACTAAAAAA 840  
QY 841 GAAGTGTCTATTAAAACTACATTAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900  
DB 841 GAAGTGTCTATTAAAACTACATTAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900

DB 841 GAAGTGTCTATTAAAACTACATTAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900  
QY 901 GACTGATGATGATGAGCAGCAGACAGTTCATATTGAATGATGGCTCAACACAGGTGGAGAA 960  
DB 901 GACTGATGATGATGAGCAGCAGACAGTTCATATTGAATGATGGCTCAACACAGGTGGAGAA 960  
QY 961 AACCTGCTGAAAAAATAAGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAAAACGCA 1020  
DB 961 AACCTGCTGAAAAAATAAGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAAAACGCA 1020  
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAACACGAGCAAACTAACCAACTTTTCACCTGCCTGAC 1080  
DB 1021 TTTGACTTAATTTAGAAAAAGCTGAACACGAGCAAACTAACCAACTTTTCACCTGCCTGAC 1080  
QY 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCAATGCTGGAATATGTTTAAAGTTTGCCTATGCT 1140  
DB 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCAATGCTGGAATATGTTTAAAGTTTGCCTATGCT 1140  
QY 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGAGCAAAAAAGAAATACGTGTTTATTTCAATGAGCA 1200  
DB 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGAGCAAAAAAGAAATACGTGTTTATTTCAATGAGCA 1200  
QY 1201 GCCAGACAGGCAAACTATTTATTCACAAGCCATAGCAACAGCACTTGGCAATGTTGCT 1260  
DB 1201 GCCAGACAGGCAAACTATTTATTTGCAAGCCATAGCAACAGCACTTGGCAATGTTGCT 1260  
QY 1261 TGTATAATGTCAGCCAACTGTAACCTTTCCATTTAATGACTGTACCAACCAAGAACTTTGATT 1320  
DB 1261 TGTATAATGTCAGCCAACTGTAACCTTTCCATTTAATGACTGTACCAACCAAGAACTTTGATT 1320  
QY 1321 TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAACACAGTTTAAAGCCATTTGCTCT 1380  
DB 1321 TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAACCAAGTTTAAAGCCATTTGCTCT 1380  
QY 1381 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGCAAAACAGAGTTGAACCAACCA 1440  
DB 1381 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGCAAAACAGAGTTGAACCAACCA 1440  
QY 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAAGAAAGACCA 1500  
DB 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAAGAAAGACCA 1500  
QY 1501 GAACACATCAACCAATCAGACAGAAATGCTTTAAACATTCATCTAAACATACCTTGCCT 1560  
DB 1501 GAACACATCAACCAATCAGACAGAAATGCTTTAAACATTCATCTAAACATACCTTGCCT 1560  
QY 1561 GGTGACTTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAG 1620  
DB 1561 GGTGACTTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAG 1620  
QY 1621 AATGGTTACCAATCTACCATGGCAAGTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATTGG 1680  
DB 1621 AATGGTTACCAATCTACCATGGCAAGTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATTGG 1680  
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAATCTCTATATAATTTACTAGGTTGGGACGC 1740  
DB 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAATCTCTATATAATTTACTAGGTTGGGACGC 1740  
QY 1741 TCACCAATTCACGACACCGGAAAGTAGCGCTCTCAGCCAGAACTATGCATTAACCTCCACTT 1800  
DB 1741 TCACCAATTCACGACACCGGAAAGTAGCGCTCTCAGCCAGAACTATGCATTAACCTCCACTT 1800  
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTGTCG 1860  
DB 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTGTCG 1860  
QY 1861 GGCATCTGCAGAAACCCAGAAACCTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920  
DB 1861 GGCATCTGCAGAAACCCAGAAACCTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920  
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGGGAAACCG 1980  
DB 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGGGAAACCG 1980

Qy	181	CGAGAGCGGAAACTTCTTGGGACCAAGCAGGACATGGAAATGGGAAACCAACAGTGGAT	240
Db	181	CGAGGAGCGGAAACTTCTTGGGACCAAGCAGGACATGGAAATGGGAAACCAACAGTGGAT	240
Qy	241	GAATGACCAAAAGCAAGTATTCTTTTGGTAAATAATGTTTAAATAATGTTTATTTGAA	300
Db	241	GAATGACCAAAAGCAAGTATTCTTTTGGTAAATAATGTTTAAATAATGTTTATTTGAA	300
Qy	301	GTGCTTTAAACAACAAGATATATTTCTCTGTGATGTTTAAATGTTTGTGCAACATGAATGG	360
Db	301	GTGCTTTAAACAACAAGATATATTTCTCTGTGATGTTTAAATGTTTGTGCAACATGAATGG	360
Qy	361	GGAAAGACCAAGCTGGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420
Db	361	GGAAAGACCAAGCTGGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420
Qy	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTCGAGCAGATGTTGTTGTAACAGCC	480
Db	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTCGAGCAGATGTTGTTGTAACAGCC	480
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
Qy	541	GAGTGGGTTTACTCTTACTTATAGCATAAGCAACCAAAAGAGACTATACCAAGTGT	600
Db	541	GAGTGGGTTTACTCTTACTTATAGCATAAGCAACCAAAAGAGACTATACCAAGTGT	600
Qy	601	GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
Db	601	GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
Qy	661	CCACCAAGAGAGCGAGGCTATTTTCTTAGCAGTACTCTGCTGGGAAATAACCTTTT	720
Db	661	CCACCAAGAGAGCGAGGCTATTTTCTTAGCAGTACTCTGCTGGGAAATAACCTTTT	720
Qy	721	AAAGAGGCGGCGGCATCTAGTGAGCAAACTATACACTGATGACATGCGCGCAGAAACG	780
Db	721	AAAGAGGCGGCGGCATCTAGTGAGCAAACTATACACTGATGACATGCGCGCAGAAACG	780
Qy	781	GTTTGAACACACAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATTAACCTAA	840
Db	781	GTTTGAACACACAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATTAACCTAA	840
Qy	841	GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCACCAG	900
Db	841	GAAAGTTCTTATTAAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCACCAG	900
Qy	901	GACTGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGCTCAACAGGTGGAGAA	960
Db	901	GACTGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGCTCAACAGGTGGAGAA	960
Qy	961	AACCTGCTGAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAGAACCAACAGCA	1020
Db	961	AACCTGCTGAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAGAACCAACAGCA	1020
Qy	1021	TTTGACTTAAATTTAGAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
Db	1021	TTTGACTTAAATTTAGAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
Qy	1081	ACAAGAACCTGCAAGATTTTGTCTTTCATGCTGGAACATATGTTTAAAGTTTGGCATGCT	1140
Db	1081	ACAAGAACCTGCAAGATTTTGTCTTTCATGCTGGAACATATGTTTAAAGTTTGGCATGCT	1140
Qy	1141	ATTGCTGTGTTTAAACAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA	1200
Db	1141	ATTGCTGTGTTTAAACAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA	1200
Qy	1201	GCCAGCACAGGCAAACTCTATTATTGCAAGCCATAGCAGCAGTGTGCAATGTTGGT	1260
Db	1201	GCCAGCACAGGCAAACTCTATTATTGCAAGCCATAGCAGCAGTGTGCAATGTTGGT	1260

Qy

1981

TTGAAGAAAGACTTTCAGGAGCGCCTGAACCTTGACTAA

2019

Db

1981

TTGAAGAAAGACTTTCAGGAGCGCCTGAACCTTGACTAA

2019

RESULT 2

AAD02797

AAD02797 standard; DNA; 2019 BP.

ID

XX

AC

DT

XX

XX

XX

XX

DE

XX

KW

KW

XX

OS

XX

FH

FT

FT

FT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX







QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTGG 1860  
DB 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTGG 1860  
QY 1861 GGCACCTGCAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
DB 1861 GGCACCTGCAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
QY 1921 CTGAGCCCAACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGGCTTCGGTGGGAACCG 1980  
DB 1921 CTGAGCCCAACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGGCTTCGGTGGGAACCG 1980  
QY 1981 TTCAAGAAAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2019  
DB 1981 TTCAAGAAAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2019

## RESULT 5

AAD02805  
ID AAD02805 standard; DNA; 2019 BP.

XX AC AAD02805;

XX DT 06-AUG-2003 (revised)

XX DT 31-MAY-2001 (first entry)

XX DE Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX KW NS1; non-structure protein 1; cytosolic; gene therapy; toxin; therapy;

XX KW tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX OS Parvovirus.

XX OS Synthetic.

XX FH Key

XX FT CDS

XX FT 1..2019

XX FT /\*tag= a

XX FT /product= "NS1 variant (T463A) protein"

XX FT replace(1387, A)

XX FT /\*tag= b

XX FT mutation

XX FT EP1077260-A1.

XX PN 21-FEB-2001.

XX PP 13-AUG-1999; 99EP-00115161.

XX PP 13-AUG-1999; 99EP-00115161.

XX PP (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX PI Nueesch J, Rommelaere J;

XX DR WPI; 2001-212717/22.

XX DR P-PSDB; AAY72710.

XX XX Novel parvovirus non-structure protein variant, useful for treating

XX PT tumoral diseases, has a shifted equilibrium between DNA replication and

XX PT transcription activities, and cytotoxic activity.

XX XX Claim 7; Page 27-30; 4lpp; English.

XX CC The present sequence is a DNA encoding parvovirus non-structure protein 1

XX CC (NS1) variant (T463A). The invention relates to the variants of the

XX CC parvovirus non-structure protein (NS1) having a shifted equilibrium

XX CC between the DNA replication and transcription activities, and the

XX CC cytotoxicity activity. These variants are useful as toxins for treating

XX CC tumoral diseases. The variant DNAs are useful as vectors for gene

XX CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX PS Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

XX XX Query Match 99.8%; Score 2015.8; DB 5; Length 2019;

Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60  
DB 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60  
QY 61 AGTAACAGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120  
DB 61 AGTAACAGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGACGAGCTGGAATCTTTACAA 180  
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGACGAGCTGGAATCTTTACAA 180  
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGACATGGAATGGGAAACACAGTGGAT 240  
DB 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGACATGGAATGGGAAACACAGTGGAT 240  
QY 241 GAAATGACCAAAAAAGCAAGTATTTCAATTTTGGTAAATGTTTAAAAATGTTTATTGAA 300  
DB 241 GAAATGACCAAAAAAGCAAGTATTTCAATTTTGGTAAATGTTTAAAAATGTTTATTGAA 300  
QY 301 GTGCTTAAACACAAAGATATATTTCTCTGGTGAATTTAAATGGTTTGGCAACATGAATGG 360  
DB 301 GTGCTTAAACACAAAGATATATTTCTCTGGTGAATTTAAATGGTTTGGCAACATGAATGG 360  
QY 361 GGAAGAGACCAAGCGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTTAGTCAAGCT 420  
DB 361 GGAAGAGACCAAGCGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTTAGTCAAGCT 420  
QY 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC 480  
DB 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC 480  
QY 481 TGTAAATGTCNACTACACAGCTGAAGAAATTTAAACTAAGAGAAATAGCAGAGACAAT 540  
DB 481 TGTAAATGTCNACTACACAGCTGAAGAAATTTAAACTAAGAGAAATAGCAGAGACAAT 540  
QY 541 GAGTGGTTTACTCTTACTTACTTATAAGCATAAGCAACCAAAAGACTATACCAAGTGT 600  
DB 541 GAGTGGTTTACTCTTACTTACTTATAAGCATAAGCAACCAAAAGACTATACCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATTTCTTACTTATTTTAACTAAAAAGAAAAATAGCACTAGT 660  
DB 601 GTTCTTTTGGAAACATGATTTCTTACTTATTTTAACTAAAAAGAAAAATAGCACTAGT 660  
QY 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACCTAATCTTTTAA 720  
DB 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACCTAATCTTTTAA 720  
QY 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780  
DB 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780  
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTAAGCGGCGCAGAAATTCAAACTAAAAA 840  
DB 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTAAGCGGCGCAGAAATTCAAACTAAAAA 840  
QY 841 GAAGTTGCTTATTAATACTACATTAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900  
DB 841 GAAGTTGCTTATTAATACTACATTAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900  
QY 901 GACTGGATGATGATGAGCAGCAGTATACATTGAATGATGCTCAACCCAGGTGGAGAA 960  
DB 901 GACTGGATGATGATGAGCAGCAGTATACATTGAATGATGCTCAACCCAGGTGGAGAA 960  
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACAACTTCTAGCCAGAACCAAAACAGCA 1020  
DB 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACAACTTCTAGCCAGAACCAAAACAGCA 1020  
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAACCCAGCAAACTAACCACTTTTTCACCTCTGAC 1080

Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAACCCAGCAAACTAACCACTTTTCACTGCGCTGAC 1080  
Qy 1081 ACAAGAACCTGCAGAAATTTTGGCTTTTATGCTGGAATCTATGTTAAAGTTTGGCCATGCT 1140  
Db 1081 ACAAGAACCTGCAGAAATTTTGGCTTTTATGCTGGAATCTATGTTAAAGTTTGGCCATGCT 1140  
Qy 1141 ATTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATCTGTTTATTTATGAGACCA 1200  
Db 1141 ATTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATCTGTTTATTTATGAGACCA 1200  
Qy 1201 GCAGACAGGGAATCTATTATGACAGCCATAGCAGCAAGCACTTGGCATGTTGCT 1260  
Db 1201 GCAGACAGGGAATCTATTATGACAGCCATAGCAGCAAGCACTTGGCATGTTGCT 1260  
Qy 1261 TGCTATAATGACGCAATGTAATCTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
Db 1261 TGCTATAATGACGCAATGTAATCTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
Qy 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT 1380  
Db 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT 1380  
Qy 1381 GGTCAAACTATTGCGATTGATCAAAAGAGGAGCAAGCAAGTAAAGTAAAGCCATTTGCTCT 1440  
Db 1381 GGTCAAACTATTGCGATTGATCAAAAGAGGAGCAAGCAAGTAAAGTAAAGCCATTTGCTCT 1440  
Qy 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1500  
Db 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1500  
Qy 1501 GAACACACTCAACCAATCAGAGACAGAACTGTTAACTCACTCAACATCACTCACTCACTCACT 1560  
Db 1501 GAACACACTCAACCAATCAGAGACAGAACTGTTAACTCACTCAACATCACTCACTCACTCACT 1560  
Qy 1561 GGTGACTTTGGTGTGACAAATGATGCGCAATGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTG 1620  
Db 1561 GGTGACTTTGGTGTGACAAATGATGCGCAATGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTG 1620  
Qy 1621 AATGGTTTACCAATCTACCTGCGAAGCTACTGTGCTAAATGGGCAAAAGTTCTGATTGG 1680  
Db 1621 AATGGTTTACCAATCTACCTGCGAAGCTACTGTGCTAAATGGGCAAAAGTTCTGATTGG 1680  
Qy 1681 TCAGAAACTGGGCGAGCCAAAGGTGCCAATCCTTATAAATTTACTAGTTCGGGACGC 1740  
Db 1681 TCAGAAACTGGGCGAGCCAAAGGTGCCAATCCTTATAAATTTACTAGTTCGGGACGC 1740  
Qy 1741 TCACCATTCAGACACCGAAGTACGCTCTCAGCCAGAACTATGCACTAATCTCCACTT 1800  
Db 1741 TCACCATTCAGACACCGAAGTACGCTCTCAGCCAGAACTATGCACTAATCTCCACTT 1800  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATCTCTGTTGG 1860  
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATCTCTGTTGG 1860  
Qy 1861 GGCACTCGAAGACCCAGAACACTGGGGAAAGTGGTTCCAAAGCTTCCAAAGATGTTCAA 1920  
Db 1861 GGCACTCGAAGACCCAGAACACTGGGGAAAGTGGTTCCAAAGCTTCCAAAGATGTTCAA 1920  
Qy 1921 CTGAGCCCAATGGTTCAGAGATCGAGGAGATTTGAGAGCGTGTCTTCGGTGGGAAACCG 1980  
Db 1921 CTGAGCCCAATGGTTCAGAGATCGAGGAGATTTGAGAGCGTGTCTTCGGTGGGAAACCG 1980  
Qy 1981 TTGAAGAAAGACTTTCAGGAGCCGCTGAATCTGGACTAA 2019  
Db 1981 TTGAAGAAAGACTTTCAGGAGCCGCTGAATCTGGACTAA 2019

RESULT 6  
AAD02803  
ID AAD02803 standard; DNA; 2019 BP.  
XX  
AC AAD02803;  
XX

DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA.  
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.  
KW Parvovirus.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH CDS 1..2019  
FT /\*tag= a  
FT /product= "NS1 variant (T394A) protein"  
FT replace(1180, A)  
FT mutation /\*tag= b  
XX  
PN EP1077260-A1.  
XX  
PD 21-FEB-2001.  
XX  
PP 13-AUG-1999; 99EP-00115161.  
XX  
PR 13-AUG-1999; 99EP-00115161.  
XX  
PA (DEK-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
XX Nueesch J, Rommelaere J;  
XX  
XX WPI; 2001-212717/22.  
DR P-PSDB; AAY72708.  
XX  
XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.  
XX  
PS Claim 7; Page 22-24; 4lpp; English.  
XX  
CC The present sequence is a DNA encoding parvovirus non-structure protein 1  
CC (NS1) variant (T394A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoral diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;  
Query Match 99.8%; Score 2015.8; DB 5; Length 2019;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAAA 60  
Db 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAAA 60  
Qy 61 AGTAACCAAGCAAGTGTCTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
Db 61 AGTAACCAAGCAAGTGTCTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
Qy 121 GATATCGGATGGATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180  
Db 121 GATATCGGATGGATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180  
Qy 181 CGAGGAGCGGAACTACTTGGGACCAAAAGCGAGCATGGAATGGGAAACCAAGTGGAT 240  
Db 181 CGAGGAGCGGAACTACTTGGGACCAAAAGCGAGCATGGAATGGGAAACCAAGTGGAT 240  
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGGTTTAAATGTTTATTGAA 300  
Db 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGGTTTAAATGTTTATTGAA 300



QY 301 GTGCTTAACACAAAGAAATATATTTCTCTGGTGATGTTAAATGGTTTGTGCAACATGAATGG 360  
Db 301 GTGCTTAACACAAAGAAATATATTTCTCTGGTGATGTTAAATGGTTTGTGCAACATGAATGG 360  
QY 361 GGAAGAAGACCAAGGCTGGCACTGCGCATGTACTTAATTCGAGGAAGAGACTTTTAGTCAAGCT 420  
Db 361 GGAAGAAGACCAAGGCTGGCACTGCGCATGTACTTAATTCGAGGAAGAGACTTTTAGTCAAGCT 420  
QY 421 CAAGGGAAATGGTGGAGAAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAACAGCC 480  
Db 421 CAAGGGAAATGGTGGAGAAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAACAGCC 480  
QY 481 TGTAAATGTGCAACTTAACACAGCTGGAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
Db 481 TGTAAATGTGCAACTTAACACAGCTGGAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
QY 541 GAGTGGGTACTCTACTTACTTAAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600  
Db 541 GAGTGGGTACTCTACTTACTTAAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATGTTACTATTTTAACTAAAGAAAGAAATTAAGCACTAGT 660  
Db 601 GTTCTTTTGGAAACATGATGTTACTATTTTAACTAAAGAAAGAAATTAAGCACTAGT 660  
QY 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAGAACTAATCTTTTA 720  
Db 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAGAACTAATCTTTTA 720  
QY 721 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
Db 721 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
QY 781 GTTGAACACACAGTAACCACTCGGCAGGAACTAAGCGCGGAGAAATTCAAACTTAAAAAA 840  
Db 781 GTTGAACACACAGTAACCACTCGGCAGGAACTAAGCGCGGAGAAATTCAAACTTAAAAAA 840  
QY 841 GAAGTGTCTATTAAACTACACTTAAGAGCTGGTGCATTAAGAGAGTAACCTACACAGAG 900  
Db 841 GAAGTGTCTATTAAACTACACTTAAGAGCTGGTGCATTAAGAGAGTAACCTACACAGAG 900  
QY 901 GACTGGATGATGACGAGCAGACAGTTTACATTGAATGATGCTCAACACAGGTGGAGAA 960  
Db 901 GACTGGATGATGACGAGCAGACAGTTTACATTGAATGATGCTCAACACAGGTGGAGAA 960  
QY 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
Db 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAAACGAGCAAACTAACCTTTTCACTGCCTGAC 1080  
Db 1021 TTTGACTTAATTTAGAAAAAGCTGAAACGAGCAAACTAACCTTTTCACTGCCTGAC 1080  
QY 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCTAGCT 1140  
Db 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCTAGCT 1140  
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACTGTTTATTTTATTCATGGACCA 1200  
Db 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTATTCATGGACCA 1200  
QY 1201 GCAGACACAGGCAAACTATTATTGCAAGGCCATAGCAAAAGAGTGTGGCAATGTTGGT 1260  
Db 1201 GCAGACACAGGCAAACTATTATTGCAAGGCCATAGCAAAAGAGTGTGGCAATGTTGGT 1260  
QY 1261 TGCTATTAATCAGCCATGTAACTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
Db 1261 TGCTATTAATCAGCCATGTAACTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
QY 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380  
Db 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380  
QY 1381 GGTCAAACTATTCCGATTGATCAAAAAGGAAAGGCGAGCAACAGATTTGAACCAACACCA 1440

Db 1381 GGTCAAACTATTCCGATTGATCAAAAAGGAAAGGCGAGCAACAGATTTGAACCAACACCA 1440  
QY 1441 GTCATCATGACCAAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGACCA 1500  
Db 1441 GTCATCATGACCAAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGACCA 1500  
QY 1501 GAACACACTCAACCAATTCAGAGACAGAAATGCTTTAACTTCACTTAACACATACCTTGCCT 1560  
Db 1501 GAACACACTCAACCAATTCAGAGACAGAAATGCTTTAACTTCACTTAACACATACCTTGCCT 1560  
QY 1561 GGTGACTTTGGTTTGGTTGACAAAATGAATGCGCCATGATTTGCTTGGTTGGTAAAG 1620  
Db 1561 GGTGACTTTGGTTTGGTTGACAAAATGAATGCGCCATGATTTGCTTGGTTGGTAAAG 1620  
QY 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCGCAAAAGTTCTGTATTGG 1680  
Db 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCGCAAAAGTTCTGTATTGG 1680  
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCGCAACTCTCTATATAAATTTACTAGGTTGCGGACGC 1740  
Db 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCGCAACTCTCTATATAAATTTACTAGGTTGCGGACGC 1740  
QY 1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACCTAACCTCCACTT 1800  
Db 1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACCTAACCTCCACTT 1800  
QY 1801 GCATCGGATCTCAGGAGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG 1860  
Db 1801 GCATCGGATCTCAGGAGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG 1860  
QY 1861 GGCACTGCGAGAAACCCAGAACACTTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Db 1861 GGCACTGCGAGAAACCCAGAACACTTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
QY 1921 CTGAGCCCACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980  
Db 1921 CTGAGCCCACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980  
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
RESULT 7  
AAT15311  
ID AAT15311 standard; DNA; 5049 BP.  
XX AAT15311;  
XX  
DT 14-OCT-1996 (first entry)  
XX  
DE Non-attenuated canine parvovirus CPV-39 passage 5 DNA.  
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;  
KW ss.  
XX Canine parvovirus.  
OS  
XX  
PN WO9614088-A1.  
XX  
PD 17-MAY-1996.  
XX  
PF 02-NOV-1995; 95WO-US014207.  
XX  
PR 08-NOV-1994; 94US-00336345.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Parriash CR, Gruenberg A, Carmichael LE;  
XX  
XX WPI; 1996-251556/25.  
XX



Qy 1915 GGTCACTGACCCCACTTGGTCTAGAGATCGAGAGGATTGAGAGCGTCTTCGGTGGC 1974  
Db 2181 GTCAAGCGAGTCCGCGTGTCTCGAAATAGAGCAGACCTGAGAGCCATCTTTACTTCT 2240  
Qy 1975 GAACCGTTTGAAGAAGACTTCAGCGA 2000  
Db 2241 GAACAAATTGGAAGAAGATTTCGAGA 2266

## RESULT 8

AAT15312  
ID AAT15312 standard; DNA; 5049 BP.

XX AC AAT15312;

XX DT 14-OCT-1996 (first entry)

XX DE Attenuated canine parvovirus CPV-39 passage 60 DNA.

XX XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;

KW ss.

XX OS Canine parvovirus.

XX FH Key Location/Qualifiers

FT misc\_feature

59

/\*tag= a

/note= "A, C or T"

FT misc\_feature

97

/\*tag= b

/note= "A, G or T"

FT misc\_feature

4745

/\*tag= c

/note= "A G or C"

FT misc\_feature

4881

/\*tag= d

/note= "A G or T"

FT XX WO9614088-A1.

XX PN 17-MAY-1996.

XX PD 02-NOV-1995; 95WO-US014207.

XX PF 08-NOV-1994; 94US-00336345.

XX PR (CORR ) CORNELL RES FOUND INC.

XX PA Parrish CR, Gruenberg A, Carmichael LE;

XX PI WPI; 1996-251556/25.

XX DR Attenuated CPV strains contg. up to 4 mutation (s) relative to control

XX PT virus - useful as a veterinary vaccine against CPV disease in animals,

XX PT such as wild or domestic dogs.

XX XX Claim 2; Page 24-27; 42pp; English.

XX PS This viral DNA is isolated from an attenuated CPV. The DNA is preferably

XX CC derived from VBI40. The DNA is cloned into a vector which is used to

XX CC transfect a host cell. The vector used is preferably pGEM32 or pGEM52.

XX CC The host cells to be transfected are selected from Norden Laboratory

XX CC feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or

XX CC canine A72 cells

XX XX Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;

SQ Query Match

Best Local Similarity 51.5%; Score 1040.4; DB 2; Length 5049;

Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Qy 1 ATGCTCGAAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGTTAAAGGAAAA 60

Qy

Db 273 ATGCTCGCAACAGTATATCTGAGGAAGTTATGAGGGAGTAAATGGTTAAAGAAACAT 332  
Qy 61 AGTAAACGAGGAAGTGTCTCATTTGTTTTAAAAATGAAAATGTTTCAACTGAATGGAAAA 120  
Db 333 GCAGAAAAATGAAGCAATTTTCGTTTGTGTTTAAATGTGCAACGTCCTCAACTAAATGGAAG 392  
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAGAGCTGAAATCTTTTACAA 180  
Db 393 GATGTTTCGCTGGAAACAACTATACCAACCAATTTCAAAATGAAGAGCTAACATCTTTAATT 452  
Qy 181 CGAGGAGCGAAACTACTCTGGGACCAAGC- - - - -GAGGACATGGAATGGGAACACCA 234  
Db 453 AGAGGAGCACAACAGCAATGGATCAACCGAAGAAGAATGGACTGGGAATCGGAA 512  
Qy 235 GTGGATGAAATGACCAAAAAAGCAAGTATTCAATTTTGTGATTTCTTTGGTTAAAAAATGTTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAAAGCAAGTACAACTTTTGTGATGCAATTAATTAATAAATGTCCT 572  
Qy 295 TTTGAAGTGTGTTAAACACAAGAATATATTTCTGCTGGTGTATTTAATTTGGTTTGTGCAACAT 354  
Db 573 TTTGAAGTCTTTGTTTCTAAAAATATAGAACCAATGATGTGTGTTGTTTATTTCAACAT 632  
Qy 355 GAATGGGAAAGACCAAGGCTGCGCACTGCTACTAATTTGGAGGAAAGGACTTTAGT 414  
Db 633 GAATGGGAAAGATCAAGGCTGGCATTTGTCATGTTTACTTTCATAGTAGAATTTACAA 692  
Qy 415 CAAGCTCAAGGAAATGGTGGAGAAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTA 474  
Db 693 CAAGCAACTGGTAAATGGCTACGACACAATGAATATGTTATTTGGAGTAGATGTTGGTG 752  
Qy 475 ACAGCTGTAAATGTGCAACTAAACACAGCTGAAAGAATTTAACTAGAGAGAAATAGCAGAA 534  
Db 753 ACTCTTTGTTGCGTAAACTTTAACCAACTGAAAGAATTTAAGCTCAGAGAAATTCAGAA 812  
Qy 535 GACAATGAGTGGGTACTCTACTTACTTATAAGCATAAGCAACCAACCAAGAGACTATACC 594  
Db 813 GATAGTGAATGGGTGACTATTAAACATACAGACATAAGCAACCAACCAAGAGACTATGTT 872  
Qy 595 AAGTGTGTTCTTTTGGAAACATGATTTGCTTACTATTATTTTAACTAAAAAGAAATTAAGC 654  
Db 873 AAAATGTTTCAATTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAATTTGTC 932  
Qy 655 ACTAGTCCCAAGAGAGAGAGGCTATTTCCTTAGCAGTGAATCTCTGCTGGAATACTAAC 714  
Db 933 CACAT- - -GACAAAAAGAAAGTGGCTATTTTTTAACTACTGATTTCTGCTGGAATTTAAC 989  
Qy 715 TTTTAAAGAGAGCGCGCCATCTAGTGAGCAAACTATACACTGATGATGATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACAATTTTACACTGAACAAATGAACCA 1049  
Qy 775 GAAACGGTTGAAACACAGTAACCACTGCGCAGGAAACTAAGCGCGGAGAAATTTCAAACT 834  
Db 1050 GAAACCGTTGAAACACAGTGAAGCAGCAGCAAGCAAGCAAGCGCGGAGAAATTTCAAACT 1109  
Qy 835 AAAAAAGAGTGTGCTATTAAAACTACACTTAAAGAGCTGGTGCATTAAGAGTAACTCA 894  
Db 1110 AAAAAAGAGTGTCAATCAATGTAATTTCTGCGGCACTTGTGTTAGTAAAGAGTAACATCA 1169  
Qy 895 CCAGAGACTGGATGATGATGAGCAGCAGAGTTCATTTGAAATGATGGCTCAACAGGT 954  
Db 1170 CTTGAAGACTGGATGATGTTTACCAACAGATAGTTATTGAAATGATGGCAGCAACAGGA 1229  
Qy 955 CGAGAAAACTGCTGAAAAAATACGCTAGAGATTTGTACACTTAATCTTAGCAGAACCAAA 1014  
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTTGTACTTTGACTTTAGCAGACAA 1289  
Qy 1015 ACAGCATTTGACTTAATTTTAAAAAGCTGAAACCGAGCAAACTAACCAACTTTTCACTG 1074  
Db 1290 ACAGCATTTGAAATTAATTAATTTGAAAAAGCAGATAATACTAACTAACTTTTGTCTT 1349  
Qy 1075 CTTGACACAAGAACTGCGAGAAATTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTGC 1134

Db	1350	GCAAAATTTAGACATGTCAAATTTT	TAGATGTCACGGATGGAATTTGGATTAAAGTTGT	1409
Qy	1135	CATGCTATTGCTGCTTTTAAACACAGCAAGGAGGCAAAAGAAATACTGTTTATTTCAT		1194
Db	1410	CAGCCTATAGCATGTGTTTAAATAGACAGGTGGTAAAGAAATACAGTCTCTTTTCAT		1469
Qy	1195	GGACCAAGCAGCAGCGCAAAATCTATTATTGCAAGCCATAGCAACAAGCTGTGCAAT		1254
Db	1470	GGACCAAGCAAGTACAGAAATCTATCAITGCTCAAGCCATAGCAACAAGCTGTGGTAAT		1529
Qy	1255	GTTGGTGTCTAATGATGAGCCAAATGAAACTTTCCATTTTAATGACGTGACCAACAGAC		1314
Db	1530	GTTGGTGTCTAATGATGAGCCAAATGAAACTTTCCATTTTAATGACGTGACCAATAAAAAT		1589
Qy	1315	TTGATTTGGGTAGAGAGCTGTAATTTGGACAGCAAGTAAACAGTTTAAAGCAAT		1374
Db	1590	TTAATTTGGATGGAAGCTGTAATTTGGTCAACAGTTTAAAGCAAT		1649
Qy	1375	TGCTCTGGTCAAACTATTTCGATTGATCAAAAAGGAAAGGCAAGCAAGATTTAAAGCAAT		1434
Db	1650	TGTTCTGGACAAACAAATTAGATTGATCAAAAAGGTAAGCAAAATTTGAACCA		1709
Qy	1435	ACACGATCATATGACCACAATAGAACATTTACAGTGTGAGATAGCTGCGAGAA		1494
Db	1710	ACTCCAGTAAATTATGACAACTAATGAAATATTAACAATTTGTGAGAAATGGATGTGAAGAA		1769
Qy	1495	AGACCAAGACACTCAACCAATCAGACAGAGATGCTTAAACATTCATCTAACACATACC		1554
Db	1770	AGACCTGAACATACACACCATAAGACAGAGATGTTGACATTAAGTTAGTATGTAAG		1829
Qy	1555	TTGCTGTGTAATTTGGTTGGTTGACAAAATGAAATGGCCCATGATTTGTGCTGGTTG		1614
Db	1830	CTTCCAGGAGACTTTGGTTGGTTGATATAAGAAAGAAATGGCTTTAATATGTGATGCTTA		1889
Qy	1615	GTAAGAATGTTACCAATCTACATGCGCAAGTACTGTCTAAATGGGGCAAGTTTCT		1674
Db	1890	GTTAAACATGTTTGGTTGATCAACCATGGCTAACTATACATCATTTGGGGAAAGTACCA		1949
Qy	1675	GATTGGTTCAGAAACTGGGGGAGCCAAAGTGCCAACTCCTATATAATTTTACTAGTTG		1734
Db	1950	GAATGGATGAAACTGGGGGAGCCTTAAATACAGAAAGTATAAATTCACCAAGTTG		2009
Qy	1735	GCAGCTCACCATTTCAGCACCGGAAAGTACGCTCTCAGCCAGCAACTATGACACTA		1794
Db	2010	AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT		2066
Qy	1795	CCATTGATCGATCTCGAGGACCTGCTTTAGAGCCTTTGGAGCACACCAAAATCTCCT		1854
Db	2067	CCTCTGACTCGGAGCGTAGTGGACCTTTGCACTGGAACCGTGGAGTACTCCAGATACGCT		2126
Qy	1855	GTTGCGGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGAT		1914
Db	2127	ATTGAGAACTGCAATCAACATCAACCAACTTGGCGTTACTCA-----CAAAGAC		2180
Qy	1915	GGTCAACTGAGCCCAACTTGTGTCAGATGATGAGAGATTTGAGAGCGTCTTCGGTGG		1974
Db	2181	GTGCAAGCGAGTCCGACGTGTGTCGGAATAGAGGACAGCCTGAGAGCCATCTTCTTCT		2240
Qy	1975	GAAACGTTGAGAAGACTTTCAGCGA	2000	
Db	2241	GAACAATTTGGAAGAGATTTTCGAGA	2266	
RESULT 9				
AAAT88324	ID	AAAT88324	standard; DNA; 5049 BP.	
XX	AC	AAAT88324;		
XX	DT	17-OCT-2003	(revised)	
XX	DT	21-MAY-1998	(first entry)	
XX	DE	Attenuated canine parvovirus (vBI440) genomic DNA.		

Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;  
feline panleukopenia virus; mink enteritis virus; infection; ds.  
Canine parvovirus; vBI440 (ATCC VR 2489).

Key

Location/Qualifiers

59

mutation

/\*tag= c

/\*note= "base 59 is G in CPV-39 (passage 5)"

97

mutation

/\*tag= d

/\*note= "base 97 is C or T in CPV-39 (passage 5)"

273..2279

CDS

/\*tag= a

/\*note= "NS1/NS2 coding region"

2286..4541

CDS

/\*tag= b

/\*note= "VP1/VP2 coding region"

4745

mutation

/\*tag= e

/\*note= "base 4745 is T in CPV-39 (passage 5)"

4881

mutation

/\*tag= f

/\*note= "base 4881 is C in CPV-39 (passage 5)"

WO9742972-A1.

20-NOV-1997.

06-MAY-1997; 97WO-US0007584.

15-MAY-1996; 96US-00647655.

(CORR ) CORNELL RES FOUND INC.

Parrish CR, Carmichael LE, Gruenberg A;

WPI; 1998-008583/01.

Canine parvovirus DNA carrying specific attenuating mutation(s) - used as  
vaccines for protection against parvovirus and feline pan-leukopenia  
virus infections.

Example 8; Page 34-37; 60pp; English.

This DNA sequence comprises an attenuated virus genome derived by serial  
passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate  
39 in NL6K feline kidney host cells. The attenuated virus is designated  
vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence  
(see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations  
are within the hairpin formed by the 3' terminal palindromic: the mutation  
at nucleotide 59 introduces an A into a G-C rich region within the tip of  
the hairpin, disrupting the base pairing in one of the 2 small internal  
palindromes within that sequence; the thymine at nucleotide 97 is  
adjacent to the mismatched bubble (flip-flop) sequence within the  
palindrome. The DNA from attenuated CPV strains (see also AAT88321) is  
used for the production of infectious molecular DNA clones, which, in  
turn, can be transfected into cells to generate master stocks of the  
virus. The attenuated viruses can be used in dogs as a vaccine to protect  
against CPV disease, or more generally in cats and minks to protect  
against feline panleukopenia virus and mink enteritis virus. The vaccines  
protect against the currently prevalent CPV-2b type (and all extant  
strains of types 2 and 2a), providing a long term immune response.  
(Updated on 17-OCT-2003 to standardise OS field)

Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

Query Match

51.5%;

Score 1040.4;

DB 2;

Length 5049;

Best Local Similarity

71.4%;

Pred. No. 4.6e-268;

Matches 1432;

Conservative 0;

Mismatches 556;

Indels 18;

Gaps 4;

1

ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA

60

Db 273 ATGCTGGCAACGATATCTGAGAAAGTTATGGAGGAGTAATGGTTAAAGAACAT 332  
Qy 61 AGTAACAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAAAA 120  
Db 333 GCAGAAAATGAAGCATTTTCGTTGTTTTTAAATGTGACAACTGCAACTAAATGGAAAG 392  
Qy 121 GATATCGGATGGATAGTTACAAAAGAGCTGCAGGAGCAGCTGGAATCTTTACAA 180  
Db 393 GATGTTGCGTGGAAACAACTATACAAACCAATTCAAATGAAGAGCTAACTCTTTAAT 452  
Qy 181 CGAGGACGGAACTACTCTGGGACCAAAGC-----GAGGACATGGAATGGGAACCA 234  
Db 453 AGAGGAGCACAACAGCAATGGATCAAAACCGAAGAGAGAAATGGCTGGGAATCGGA 512  
Qy 235 GTGGATGAATGACCAAAAAGCAAGTATTCATTTTGTGTTGTTGTTGTTGTTGTTGTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAGCAAGTACAAACTTTTGTGATGCAATTAATTAATAAATGCTT 572  
Qy 295 TTTGAAGTGTGTTAACACAAAGAAATATATTTCTGCTGGTGAATGTTTGGTTTGTGCAACAT 354  
Db 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAATGAATGTGTTGTTGTTTATTTCAACAT 632  
Qy 355 GAATGGGGAAGACCAAGCTGCGACTGCGATGCTACTTAATTTGGAGGAAGGACTTTAGT 414  
Db 633 GAATGGGGAAGAAATCAAGCTGCGATGTCATGTTTCTTCTCATAGTAAGAACTTACAA 692  
Qy 415 CAAGCTCAAGGGAATGGTGGAGAAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTA 474  
Db 693 CAAGCAACTGTTAAATGGCTACGACAGACAAATGAATGATGTTATGGAGTAGATGGTTGGTG 752  
Qy 475 ACAGCTGTAATGTGCAACTAACACAGCTGGAAGAAATTAACCTAGAGAGAAATAGCAGAA 534  
Db 753 ACTCTTTGTTGCGTAACTTTAACACCAACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812  
Qy 535 GACAAAGAGTGGTACTCTACTTACTTATTAAGCATAAAGCAACCAAAAGAACTATPACC 594  
Db 813 GATAGTAATGGTGGACTATTAACATACAGACATGAAGCAACAAAGAAAGAACTATGTT 872  
Qy 595 AAGTGTGTTCTTTTGGAAACATGATGTTCTTACTTATTTTAACTTAAAAAGAAATAAGC 654  
Db 873 AAAATGGTTCAATTTGGAAATATGATAGCATATTTACTTTTAAAGAAAGAAATTTGTC 932  
Qy 655 ACTAGTCCACAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGCTGGGAAACCTAAC 714  
Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGTTGGAAATTTAAC 989  
Qy 715 TTTTAAAGAGCGGAGCGCCATCTAGTCGCAAACTATACACTGATGATGATGATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGAACCA 1049  
Qy 775 GAAACGGTTGAAACCAAGTAACCACTGCGCAGGAAACTAAGCGCGGAGAAATTCAAACT 834  
Db 1050 GAAACGGTTGAAACCAAGTGAACGACAGACAGGAAACAAAGCGCGGAGAAATTCAAACT 1109  
Qy 835 AAAAAAGAGTGTCTATTTAAACCTACACTTAAAGAGCTGTGCAATAAAGAGTAACCTCA 894  
Db 1110 AAAAAAGAGTGTCAATCAAAATGTTCTTGGGAGCTTGGTTAGTAAAGAGTAACATCA 1169  
Qy 895 CCAGAGACTGGATGATGATGACGACAGACAGTTTACATTTGAATGATGGCTCAACAGGT 954  
Db 1170 CTTGAAGACTGGATGATGTTTACACCCAGATAGTTATATTGAATGATGGCAACACAGGA 1229  
Qy 955 GGAGAAACCTGCTGAAAAATACCTAGAGATTTGTACACTTAATCTAGCCAGAACCAAA 1014  
Db 1230 GGTGAAATCTTTTAAAAAATACACTTTGAATTTGTACTTTGACTTTTAGCAAGACAAA 1289  
Qy 1015 ACAGCAATTTGACTTAATTTTAGAAAAAGCTGAAACACGACAACTAACCACTTTTCACTG 1074  
Db 1290 ACAGCAATTTGAATTAATTTAGTAAAAAGCAGATAATACTAAACTAACTTTGATCTT 1349  
Qy 1075 CTTGACACAGAACTCGAATTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTGTC 1134

Db 1350 GCAAAATTTAGAACATGTCAAATTTTTTAGAATGCACGGATGGAATGGATTAAAGTTGT 1409  
Qy 1135 CATGCTATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATACTGTTTTATTTCAT 1194  
Db 1410 CACGCTATAGCATGTGTTTAAATAGACAAGGTGGTAAAAAGAAATACAGTTCTTTTTTCAT 1469  
Qy 1195 GGACCAAGCAGCAGCAGCAAACTATTTATTTGCAAGCCATAGCACAAGCAGTTGGCAAT 1254  
Db 1470 GGACCAAGCAGCAGCAGCAAACTATTTATTTGCTCAAGCCATAGCACAAGCTGTGGGTAAT 1529  
Qy 1255 GTTGGTTGCTATTAATCAGCCAAATGTAAACTTTTCCATTTTAATGACTGTACCAACAAGAAC 1314  
Db 1530 GTTGGTTGTTATAATCAGCAAAATGTAAATTTTTCCATTTTAATGACTGTACCAATAAAAT 1589  
Qy 1315 TTGATTTGGGTAGAAAGCTGTAACTTTTGGACAGCAAGTAAACCGATTTTAAAGCCATT 1374  
Db 1590 TTAATTTGGATTGAAGAAGCTGTAACTTTTGGTCAACAGTTAATCAATTTTAAAGCAATC 1649  
Qy 1375 TGCTCTGGTCAAACTATTCGCTTGTATCAAAAGGAAAGGAGCAACAGATTTGAACCA 1434  
Db 1650 TGTCTGGACAAACAAATTTGAATTTGATCAAAAGGTAAGGAAATGAACCA 1709  
Qy 1435 ACACCACTCATCATGACCAAAATGAGAACATTTACAGTGTGTCAGATAGGCTCGGAAGAA 1494  
Db 1710 ACTCCAGTAATATGACCAACTAATGAATATATACAAATTTGTGAGAAATTTGAAGAA 1769  
Qy 1495 AGACCAAGACACACTCAACCAATCAGACAGAGAAATGCTTAACTTAACTTAAACATACC 1554  
Db 1770 AGACCTGAACATACACAAACCAATGAAGAGACAGAAATGTGAACATTAAGTTAGTATGAAG 1829  
Qy 1555 TTGCTCTGGTGAATTTGGTTGTTGACAAAATGAATGGCCCATGATTTGCTGTGGTTG 1614  
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAGAAAGAAATGGCCTTTAATATGTGATGGTTA 1889  
Qy 1615 GTAAAGAAATGGTTACCAATCTACATGCGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCCCT 1674  
Db 1890 GTTAAACATGGTTTGTGATCAACCATGGCTTAACATACACATCATTTGGGGAAGTACCA 1949  
Qy 1675 GATGCTGAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCTTATATAATTTTACAGTTGCG 1734  
Db 1950 GAATGGATGAAAACCTGGGCGGAGCCTTAAATACAAAGAAAGTATAAATTCACCAGGTTCG 2009  
Qy 1735 GCAGCTCACCATTACGACACCGAAAGTACGCTCTCAGCCAGACACTATGCATTAAC 1794  
Db 2010 AAAGACT---TAGAGACAAAGCGCAAGCAACTCTCAGAGTCAAGACCAAGTTCTAACT 2066  
Qy 1795 CCACTTGCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAAAATACCTCT 1854  
Db 2067 CTTCTGACTCCGAGCGTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCT 2126  
Qy 1855 GTTGGCGGCACTGCAGAAAACCCAGAACACTGGGGAAAGCTGGTTCCAAAGCCTGCAAGAT 1914  
Db 2127 ATTGCAGAACTGCAAAATCAACAACTCAACCAACTTTGGCGTTACTCA-----CAAGAC 2180  
Qy 1915 GGTCAACTGAGCCCACTTGGTCAGAGATCGAGAGGATTTGAGAGGCTGCTTCGGTGGC 1974  
Db 2181 GTGCAAGCGAGTCCGAGCTGGTCCGAAATAGAGAGCAGACCTGAGAGCCTCTTACTTCT 2240  
Qy 1975 GAACCGTTGAAGAAAGAACTTCAGCGA 2000  
Db 2241 GAACAATTTGGAAGAAAGATTTTCAGA 2266

RESULT 10

AAT88321

ID AAT88321 standard; DNA; 5049 BP.

XX AAT88321;

AC AAT88321;

XX 21-MAY-1998 (first entry)

XX Attenuated canine parvovirus genomic DNA.

XX



```
Db 1230 GGTGAAATCTTTAAAAATACACTTGAAATTTGTACTTTGACTTTAGCAAGACAAAA 1289
Qy 1015 ACAGATTGACTTAATTTTAGAAAAAGCTGAACACGACAAACTPAACAACTTTTCACTG 1074
Db 1290 ACAGCAATTGAATTAATCTTGAAAAAGCAGATAATACTAAACTTAACTAACTTTGATCTT 1349
Qy 1075 CCTGACACAGAACTCGCAGAAATTTTTCCTTTTCATCGCTGGAACATATGTTAAAGTTTGC 1134
Db 1350 GCAAATCTAGACATGTCAAATTTTGTAGATGCACGGATGGAATTTGGATTTAAAGTTTGT 1409
Qy 1135 CATGCTATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTTTATTTTCAT 1194
Db 1410 CACGCTATGATGTGTTTTTAAATAGACAAAGTGTAAAGAAATACAGTCTTTTTCAT 1469
Qy 1195 GGACGACGACGACAGGCAAACTTATTATTGCAACAAGCCATAGACAAAGCAGTTGGCAAT 1254
Db 1470 GGACGACGACGACGAGAAATCTATCATTTGCTCAAGCCATAGACAAAGCTGTGGGTAAT 1529
Qy 1255 GTTGGTTGCTATATGCTGAGCCAAATGTAACCTTTTCCATTTAATGACTGTACCAACAAGAC 1314
Db 1530 GTTGGTTGTTATATAGCAGCAATGTAAATTTTCCATTTAATGACTGTACCAATAAAAT 1589
Qy 1315 TTGATTTGGGTAGAAAGCTGTGTAACCTTTGGACAGCAAGTAACCACTTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATTGAAGAGCTGTGTAACCTTTGGTCAACAAGTTAATCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTGCAATTTGATCAAAAAGGAAAGGCAAGCAACAGATTGAACCA 1434
Db 1650 TGTCTGCAACAACTAGAAATTTGATCAAAAAGGTAAGCAAGTAAAGCAAAATGAACCA 1709
Qy 1435 ACACCACTCATGACCACAAAATGAAACATTTACAGTGTGTCAGAAATAGGCTGCGAAGAA 1494
Db 1710 ACTCCAGTAATATTGACAACTAATGAAATTTTAACAATTTGGAATTTGATGTGAAGAA 1769
Qy 1495 AGACCAACACACTCAACCAATCAGACAGAGATGCTTAAACATTTTCACTTAAACACATACC 1554
Db 1770 AGACCTGACATACACAAACCAATAGACAGACAGATGTTGAACATTAAGTAGTAGTAGAG 1829
Qy 1555 TTGCTGGTGAATTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTGTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAAGATGGCTTTTAAATGTGCGATGTTA 1889
Qy 1615 GTAAAGAAATGGTTACCAATCTACATGCGAAGCTACTGTGCTAAATGGGCAAAAGTTCTT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGCTCACTATACATCATCTGGGAAAAAGTACCA 1949
Qy 1675 GATTGGTCAGAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTTACTAGGTTCG 1734
Db 1950 GAATGGATGAAACTGGCGGAGCCCTAAATACAGAAAGTATAAATTCACCGGTTGC 2009
Qy 1735 GCAGCTCACCAATTCAGCACACGAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCATTGCTCGATCTGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTT 1854
Db 2067 CTTCTGACTCCGGACGTAGTGGACCTTGGCACTGGAACCCGTGGAGTACTCCAGATACGCCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCTGCGCAAGAT 1914
Db 2127 ATTGCAGAACTGCAATTAACAACTCAACCACTTGGCGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGC 1974
Db 2181 GTGCAAGGAGTCCGACGTGTGTCGNAATAGAGGACACCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAGACTTCAGCGA 2000
Db 2241 GAACAAATTGGAAGAAGATTTTCGAGA 2266
```

RESULT 11  
AAT88320

```
ID XX AAT88320 standard; DNA; 5049 BP.
XX AC AAT88320;
XX XX 17-OCT-2003 (revised)
DT 21-MAY-1998 (first entry)
XX XX Canine parvovirus 39 passage #5 (wild-type).
DE Canine parvovirus; CPV; attenuation; vB1440; vaccine; dog;
XX Canine panleukopenia virus; mink enteritis virus; infection; ds.
XX Canine parvovirus; type 2b isolate 39.
XX Key Location/Qualifiers
FH CDS 273..2279
FT FT /*tag= a
FT FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT FT /*tag= b
FT FT /note= "VP1/VP2 coding region"
XX WO9742972-A1.
XX PN 20-NOV-1997.
XX PD 06-MAY-1997; 97MO-US007584.
XX PF 15-MAY-1996; 96US-00647655.
XX PR (CORR ) CORNELL RES FOUND INC.
XX PA Parrish CR, Carmichael LE, Gruenberg A;
XX PI WP1; 1998-008583/01.
XX DR Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX PT vaccines for protection against parvovirus and feline pan-leukopenia
XX PT virus infections.
XX PS Example 8; Page 37-40; 60pp; English.
XX CC This DNA sequence comprises the genome of virulent canine parvovirus type
CC 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NUKF feline
CC kidney host cells. Further passaging has yielded attenuated virus vB1440
CC (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
CC (see AAT88321) derived from the 65th passage (ATCC 2528). These
CC respectively contain 4 and 6 mutations in comparison to the virulent 5th
CC passage virus. The DNA from attenuated CPV-2b strains is used for the
CC production of infectious molecular DNA clones, which, in turn, can be
CC transfectd into cells to generate master stocks of the virus. The
CC attenuated viruses can be used in dogs as a vaccine to protect against
CC CPV disease, or more generally in cats and minks to protect against
CC feline panleukopenia virus and mink enteritis virus. The vaccines protect
CC against the currently prevalent CPV-2b type (and all extant strains of
CC types 2 and 2a), and provide a long term immune response. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;
Best Local Similarity 71.3%; Pred. No. 1.2e-267;
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;
Qy 1 ATCGCTCGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA 60
Db 273 ATGCTCGGCAACCACTATCTAGGAGGATTTATGAGGAGTAATTTGGTTAAAGAACAT 332
Qy 61 AGTAACCAAGGAAGTGTCTCATTTGTTTGTATTTTAAATAATGAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAAAATGAAGCATTTTCGTTGTTTAAATGTGACACACGTCCAACTTAATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTACAAAAAAGAGCTGCAGAGGACGAGCTGGAATCTTTTACAA 180
```





```
FT      /*tag= a  
FT      /note= "see AAP40306"  
FT      2107. .3522  
CDS      /*tag= b  
FT      /note= "see AAP40675"  
XX  
PN      WO8402847-A.  
XX  
PD      02-AUG-1984.  
XX  
PF      19-JAN-1984; 84WO-US000063.  
XX  
PR      19-JAN-1983; 83US-00459203.  
PR      06-JAN-1984; 84US-00567968.  
XX  
PA      (AMGE-) AMGEN.  
XX  
PI      Fox GM;  
XX  
XX      WPI; 1984-201354/32.  
DR      P-PSDB; AAP40306, AAP40675.  
XX  
XX      Polypeptide obtd. by recombinant DNA methods - for vaccination against  
PT      parvovirus infections in man and animals.  
XX  
XX      Claim 10; Table II, Page 33-49; 80pp; English.  
XX  
XX      The inventors claim an immunologically active polypeptide for the  
CC      development of vaccinal immunity against parvovirus infection. Also  
CC      claimed are DNA sequences wholly or partly duplicative of defined  
CC      sequences. The polypeptides are used in vaccines for conferring  
CC      protection against parvovirus infections in man and animals. (Updated on  
XX      24-OCT-2003 to standardise OS field)  
XX  
SQ      Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;  
  
Query Match 44.7%; Score 902.6; DB 1; Length 3524;  
Best Local Similarity 69.8%; Pred. No. 3.6e-231;  
Matches 1297; Conservative 0; Mismatches 544; Indels 18; Gaps 5;  
  
QY      4 GCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAAAGT 63  
DB      13 GCGGAAACACTTACTCGGAAGAGGTACTAAAGCTTACCACTGGCTTCAAGATAATGCT 72  
QY      64 AACGAGGAAGTGTCTCATTTGTTTTTAAATAATGAAATGTTCAACTGAAATGGAAGAAT 123  
DB      73 CAAAAGAAGCAATCTCTATGATTTTAAACACAAAAAGTCAATCTAAATGGAAGAA 132  
QY      124 ATCGGATGGAATGTTACAAAAAAGAGCTGCAGAGACGAGCTGAAATCTTTACACGA 183  
DB      133 ATTGCTTGGAAATAACTACAAAGATAACACAGATGCGGAAATGATAAACCTTACAAAGA 192  
QY      184 GGAGCGAATACTTTGGGACCAAGCGAGAGATGGAATGGGAAACCAACGATGGATGAA 243  
DB      193 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCGACAGC 252  
QY      244 ATGACCAAAAAGCAAGTATTCATTTTGTATTTGTTTAAATAATGTTTATTTGAAGTG 303  
DB      253 CTCACAAAAGCGCAGTACTGATTTTGTCTCTTTGTTTAAATAATGTTCTTTTGAAGGT 312  
QY      304 CTTAAACAAGAATAATATTTCTGGTGATGTTAAATGTTTGTGCAACATGAATGGGA 363  
DB      313 ATATTGCAAAAGAACCTTAAGTCCAAGTGACTGCTACTGTTCTTACAGCATGAACATGTT 372  
QY      364 AAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACCTTTAGTCAAGCTCAA 423  
DB      373 CAAGATACTGGCTATCACTGCCATGTACTAGAGTGGGAAAGGCTTTACAAAGCAATG 432  
QY      424 GGGAAATGGTGGGAAGGCAACTAAATGTTTACTGGGACAGATGGTTGGTAAACAGCCTGT 483  
DB      433 GGAATATGGTTACGAACAACTTAAACAATTTATGGATAGATGGTTGAATATCAATGC 492  
QY      484 ANTGTGCAACTAACACACGAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAAGCAATGAG 543  
  
Db      493 AAAGTACCTCTAACACAGTTTGAAGAATAAATAAATTAAAGGAATTTAGCAGAGGATGGTGAG 552  
QY      544 TGGGTTACTCTACTTACTTATAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGTGTT 603  
Db      553 TGGGTATCGCTACTTAACCTACACTCACAAACAACCTAAAAAACAATATACAAAAAGACT 612  
QY      604 CTTTTTGGAAAAATGATGCTTACTATCTATTTTTTAACTAAATAAAGAAAAATAGCAGCTAGTCCA 663  
Db      613 CTTTTTGGAAAAATGATGCTTACTTACTTCTTAATAAATAAAGAAAAAGACAACCT----- 666  
QY      664 CCAAGAGACGGAGGCTATTTCTTACAGTACGCTGCTGGCTGGGAAAACTACTTTTAAAAA 723  
Db      667 GAAAGAGAGCATGGATATTATCTCAGCTCAGATTTCTGGCTTCATGACAAATTTCTTAAAA 726  
QY      724 GAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACGGTT 783  
Db      727 GAAGGCGAGAGACACTTAGTCAGTCACTTATTTACTGAGCAATATAAACCTGAAACGTG 786  
QY      784 GAAACACAGTAAACCACTGCGCAGGAAACTAAGCGCGGAGAAATTCAAACTAAAAAAGAA 843  
Db      787 GAAACAAACGGTTACTACAGCTCAGGAAGTTCCCGAGGCGAGAAATACAAAACAAAAAAGAA 846  
QY      844 GTTGCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAGGAC 903  
Db      847 GTAAGCATAAATGCAATAAAGAGACTTGGTTTAAATAAAGATGTACTAGCATAGAAGGC 906  
QY      904 TGGATGATGTCAGCGCAGACAGTTCATATTGAATGATGCTCAACCCAGGTGGAGAAAAAC 963  
Db      907 TGGATGATGACAGATCCAGACAGTATATAGAAATGATGGCTCAACCCGAGGAGAAAT 966  
QY      964 CTGCTGAAAAAATACGCTAGAGATTTGTACAACTCTCTAGCCAGAACCAAAACAGCATTT 1023  
Db      967 TTAATCAAAAATACACTAGAAATAACAACCTCTTACTCTAGCAAGAAACAAAACAGCATAT 1026  
QY      1024 GACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGCACACA 1083  
Db      1027 GACTTAATTTTGAAGGCAAAACCAAGCATGCTTACCAACTTTAATATTTAGCAATACA 1086  
QY      1084 AGAACCTGCAGAAATTTTGTCTTTCATGCTGGAATCTATGTTAAAGTTTCCCATGCTATT 1143  
Db      1087 AGAACATGTAAATATTTAGCATGCACTTGGAACTACATTAAGTCTGCCATGCTATA 1146  
QY      1144 TGTGTGT-TTTAAACAGACAGGAGGCAAAAGAAATA-----CTGTTTATTTTCATGGA 1197  
Db      1147 ACTGTGTACTGAAACAGACAGGAGGAAAAAGAAATACAAATCTATTTCTATGTCATGGG 1206  
QY      1198 CAGCCAGCAGCAGGCAATCTATTATTGACAGCCATAGCAGACAGGAGTTGGCAATGTT 1257  
Db      1207 CCAGCATCAACAGGAAAAAGTAAATTTGCTCAACACATTTGCAAACTTAGTTGTAATGTT 1266  
QY      1258 GGTGTCTAATGTCAGCAATGTAAACTTTTCCATTTTAACTGCTGTACCAACAAGAACTTG 1317  
Db      1267 GGTGTCTAATGTCAGCAATGTAAACTTTTCCATTTTAACTGCTGTACCAACAAGAACTTG 1326  
QY      1318 ATTTGGGTGAGAAAGCTGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGC 1377  
Db      1327 ATATGATTTGAAGAAGCAGGAAACTTCTCTAACCAAGTAAACCAATTTCAAAGCCATATGT 1386  
QY      1378 TCTGGTCAAACTATTTGCAATTTGATCAAAAGGAAAAAGGCAAGCAACAGATTGAACCAACA 1437  
Db      1387 TCAGGTCAACCAATTTAGAAATTCACAAAAGGTAAGGAGCAACCAAAATTTGAACCAACT 1446  
QY      1438 CCACTCATCATGACCAAAATGAGAACATTTACAGTGGTTCAGATAGGCTGCGGAAGAAGA 1497  
Db      1447 CTTGTAATTAATGACTCAAAATGAAGACATACCTAAAGTTAGATAGGATGCGGAGGAAGA 1506  
QY      1498 CCAGAACACACTCAACCAATTCAGAGACAGAACTCTTAAACATTTTCACTTAACACATACCTTG 1557  
Db      1507 CCAGAACATACACAAACCAATTAAGAGACAGAACTGTTAAACATACACACCTTAACCAAGAACTG 1566  
QY      1558 CTTGGTGACTTTGTTGTTTGAACAAAATGAATGGCCCATGATTTGTGCTGCTGGTGA 1617
```

1567 CCAGGTGATTTTGGACTTTTAGAAGAACTGAATGGCCACTAATATGTGCTTGGTTGGTA 1626  
1618 AGAATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCTGAT 1677  
1627 AGAAAGGTTACCAAGCAAAATGGCTAGTATATGTCATCATTTGGGAAATGTACCTGAT 1686  
1678 TGTCTAGAAAACCTGGGGAGGCGCAAGGTGCGCAACTCTTATATAATTTACTAGTTCCGCA 1737  
1687 TGTCTAGAAAATTTGGAGGAGCCAAAATGCTATCCCAATAATACACCA--ACAGAC 1743  
1738 CGTCTACCAATTCAGCACCGCAAAAGTACGCTCTCAGCCAGACACTATGCACTAATCCA 1797  
1744 TCTCAGATTTCCACATCAGTGAATCTTCGCCAGCGGACATCACTACGCGCAACTCCA 1803  
1798 CTTCGATCGATCTCGA--GGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATATCTCC 1853  
1804 ATACAGGAGGACCTGGATTTAGCTTTTGGAGCGCTGGAGCGAGCGCAACACACC 1862

## RESULT 13

ABQ95626

ID ABQ95626 standard; DNA; 374 BP.

XX AC ABQ95626;

XX DT 28-OCT-2002 (first entry)

XX DE Tumour suppression-related oligonucleotide #1277.

XX KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
XX KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
XX KW viral infection; cell degeneration disease; neurodegeneration; ds;  
XX KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

XX OS Homo sapiens.

XX PN FR2819824-A1.

XX PD 26-JUL-2002.

XX PF 23-JAN-2001; 2001FR-00000899.

XX PR 23-JAN-2001; 2001FR-00000899.

XX PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX PI Telerman A, Anson R, Tuijnder M, Susini L;

XX DR WPI; 2002-610803/66.

XX PT New nucleic acid implicated e.g. in tumor suppression, useful for  
XX PT diagnosis of tumors, viral infection and cellular degeneration and for  
XX PT drug screening.

XX PS Claim 1; Page 360; 623pp; French.

XX CC The present invention relates to novel human nucleic acid sequences (I).  
XX CC The present sequence is one such nucleic acid sequence. Expression of (I)  
XX CC are implicated in tumour suppression or reversion and apoptosis and viral  
XX CC resistance. (I) are useful as probes or primers for detecting,  
XX CC identifying, measuring and/or amplifying nucleic acid sequences, as  
XX CC antisense reagents and for recombinant production of polypeptides. (I),  
XX CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
XX CC these vectors and antibodies (Ab) against (II) are all useful for  
XX CC treatment/prevention of viral, tumour and cell degeneration diseases  
XX CC (especially neurodegeneration, such as Alzheimer's disease and  
XX CC schizophrenia). Analysing the expression of (I) is also useful for  
XX CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
XX CC (I) are used for studying the aetiology of these diseases (also immune  
XX CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
XX CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
XX CC in the specification

SQ Sequence 374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;  
Query Match 15.2%; Score 307.4; DB 6; Length 374;  
Best Local Similarity 88.2%; Pred. No. 6.2e-72;  
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
QY 174 TTACAACAGAGGAGCGGAAACTTCTGGGACCAAGCCAGGACATGGATGGGAAACCAC 233  
DB 1 TTACAACAGAGGCGGAGACCCTTGGGACCAAGCCAGGACATGGATGGGAGCGC 60  
QY 234 AGTGGATGAATGACCAAAAGCAAGTATCTTTTGGATCTTTTGGTTAAAAAATGTT 293  
DB 61 AGTGGATGACATGACCAAAAGCAAGTATTTTATTTTGGATCTTTTGGTTAAGAGTGT 120  
QY 294 ATTTGAAGTGTCTTAACACAAAGATATATTTCTCGTGTATGTTAATGTTTGTGCAACA 353  
DB 121 GTTTGAAGTGTCTCAGCACAAAGACATAGCTCTAGTAAATGTTTACTTGGTTCGTGACGA 180  
QY 354 TGAATGGGAAAAGACCAAGGCTGCGACCTGCTACTATTAATGGAGGAAAGCACTTTAG 413  
DB 181 TGAATGGGAAAAGACCAAGGCTGCGACCTGCTCATGTCTGATTTGGAGGCAAGCACTTTAG 240  
QY 414 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTT 473  
DB 241 TCAACTCAAGGAAATGTTGGAGAGGCAAGTAAATGTTGTTACTGGAGTANATGTTG 300  
QY 474 AACAGCTGTAAATGTGCAACTAACACCAAGCTGAAAGAAATTAACAAAGAAATAGCAGA 533  
DB 301 GACTGCTGNAATGTNCAACTAACACCAAGCTGAAAGAAATTAACAAAGAAATAGCAGA 360  
QY 534 AGACATAGAGTGG 546  
DB 361 GGACACATGGANNG 373

## RESULT 14

ABQ94779

ID ABQ94779 standard; DNA; 421 BP.

XX AC ABQ94779;

XX DT 28-OCT-2002 (first entry)

XX DE Tumour suppression-related oligonucleotide #430.

XX KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
XX KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
XX KW viral infection; cell degeneration disease; neurodegeneration; ds;  
XX KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

XX OS Homo sapiens.

XX PN FR2819824-A1.

XX PD 26-JUL-2002.

XX PF 23-JAN-2001; 2001FR-00000899.

XX PR 23-JAN-2001; 2001FR-00000899.

XX PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX PI Telerman A, Anson R, Tuijnder M, Susini L;

XX DR WPI; 2002-610803/66.

XX PT New nucleic acid implicated e.g. in tumor suppression, useful for  
XX PT diagnosis of tumors, viral infection and cellular degeneration and for  
XX PT drug screening.

XX PS Claim 1; Page 141; 623pp; French.

XX CC The present invention relates to novel human nucleic acid sequences (I).

CC The present sequence is one such nucleic acid sequence. Expression of (I)  
CC are implicated in tumour suppression or reversion and apoptosis and viral  
CC resistance. (I) are useful as probes or primers for detecting,  
CC identifying, measuring and/or amplifying nucleic acid sequences, as  
CC antisense reagents and for recombinant production of polypeptides. (I),  
CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
CC these vectors and antibodies (Ab) against (II) are all useful for  
CC treatment/prevention of viral, tumour and cell degeneration diseases  
CC (especially neurodegeneration, such as Alzheimer's disease and  
CC schizophrenia). Analysing the expression of (I) is also useful for  
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
CC (I) are used for studying the aetiology of these diseases (also immune  
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
CC in the specification  
XX  
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.2%; Score 285.8; DB 6; Length 421;  
Best Local Similarity 87.6%; Pred. No. 4.1e-66;  
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
QY 173 CTTTACACGAGGAGCGAACTCTTGGACCAAGCGAGCATGGAATGGGAACCA 232  
DB 59 CTTTACACGAGGCGCGAGACCACTTGGACCAAGCGAGCATGGAATGGGAGCG 118  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTGGTTAAAAAATGTT 292  
DB 119 CAGTGGATGACATGACCAAAAGCAAGTATTCATTTTGGATTTCTTGGTTAAGAAGTGT 178  
QY 293 TATTGGAAGTCTTAACACAAAGATATATTCCTGGTGATGTTAATGTTTGTGAAC 352  
DB 179 TGTTTGAAGTGCTCAGACAAAGAACATAGCTCCTAGTAATGTTTACTTGGTTCGTGAGC 238  
QY 353 ATGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTTA 412  
DB 239 ATGAATGGGAAAGACCAAGGCTGGCACTGTCTATGTCATGTCATTTGGAGGCAAGACTTTA 298  
QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTTAAATGTTTACTTGGAGCAGATGTTGG 472  
DB 299 GTCAACCTCAAGGAAATGGTGGAGAGGCAAGTAAATGTTTACTTGGAGTAGATGTTGG 358  
QY 473 TAACAGCTCTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAAT 527  
DB 359 TGACTGCTCTAATGTGTTCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAAT 413

## RESULT 15

ABQ94724

ID ABQ94724 standard; DNA; 423 BP.

AC

XX ABQ94724;

XX

XX 28-OCT-2002 (first entry)

DT

XX Tumour suppression-related oligonucleotide #375.

DE

XX Tumour; cytostatic; antiviral; neuroprotective; neurotropic; neuroleptic;  
XX tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
XX viral infection; cell degeneration disease; neurodegeneration; ds;  
XX Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS

XX Homo sapiens.

XX

XX FR2819824-A1.

XX

XX 26-JUL-2002.

XX

XX 23-JAN-2001; 2001FR-00000899.

XX

XX 23-JAN-2001; 2001FR-00000899.

XX

XX (MOLE-) MOLECULAR ENGINES LAB SA.

PA

XX

PI

XX

XX

DR

XX

XX

PT

PT

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

SQ

Telerman A, Amson R, Tuijnder M, Susini L;

WPI; 2002-610803/66.

New nucleic acid implicated e.g. in tumor suppression, useful for  
diagnosis of tumors, viral infection and cellular degeneration and for  
drug screening.

Claim 1; Page 125-126; 623pp; French.

The present invention relates to novel human nucleic acid sequences (I).  
The present sequence is one such nucleic acid sequence. Expression of (I)  
are implicated in tumour suppression or reversion and apoptosis and viral  
resistance. (I) are useful as probes or primers for detecting,  
identifying, measuring and/or amplifying nucleic acid sequences, as  
antisense reagents and for recombinant production of polypeptides. (I),  
polypeptides (II) encoded by (I), vector containing (I), cells containing  
these vectors and antibodies (Ab) against (II) are all useful for  
treatment/prevention of viral, tumour and cell degeneration diseases  
(especially neurodegeneration, such as Alzheimer's disease and  
schizophrenia). Analysing the expression of (I) is also useful for  
diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
(I) are used for studying the aetiology of these diseases (also immune  
and inflammatory diseases). Note: In the present specification, SEQ ID 1  
to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
in the specification

Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.1%; Score 284.8; DB 6; Length 423;  
Best Local Similarity 89.2%; Pred. No. 7.6e-66;

Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGAACTCTTGGACCAAGCGAGCATGGAATGGGAACCA 232

DB 71 CTTTACACGAGGCGCGAGACCACTTGGACCAAGCGAGCATGGAATGGGAGCG 130

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTAAAAAATGTT 292

DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTAAGAAGTGT 190

QY 293 TATTGGAAGTCTTAACACAAAGATATATTTCTCGTGATGTTAATGTTTGTGCAAC 352

DB 191 TGTTTGAAGTGCTCAGCACAAAGACATAGCTCCTAGTATGTTTACTTGGTTCGTGAGC 250

QY 353 ATGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTTA 412

DB 251 ATGAATGGGAAAGGACCAAGGCTGGCACTGTCTATGCTGATTTGGAGGCAAGACTTTTA 310

QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTTAAATGTTTACTTGGAGCAGATGTTGG 472

DB 311 GTCAACCTCAAGGAAATGGTGGAGAGGCAAGTAAATGTTTACTTGGAGTAGATGTTGG 370

QY 473 TAACAGCTCTAATGTGCAACTAACACAGCTGAAAGAAATTAATAA 516

DB 371 TGACTGCTCTAATGTTTCAACTAACACAGCTGAAAGAAATTAATAA 414

Search completed: January 22, 2005, 11:49:41

Job time : 958.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds  
(without alignments)  
8723.907 Million cell updates/sec

Title: US-10-069-056-4

Perfect score: 2019

Sequence: 1 atggctggaactgttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	109.4	5.4	4680	1	US-08-254-358-1
6	109.4	5.4	4680	1	US-08-475-391-1
7	109.4	5.4	4680	2	US-08-709-609-1
8	109.4	5.4	4680	5	PCT-US95-07178-1
9	109.4	5.4	4681	4	US-09-807-802A-18
10	109.4	5.4	4910	2	US-08-331-384-2
11	109.4	5.4	4910	3	US-08-836-087-2
12	109.4	5.4	4910	3	US-09-246-320-2
13	109.4	5.4	4910	3	US-09-546-738-2
14	109.4	5.4	7214	4	US-09-438-268-1
15	109.4	5.4	7557	4	US-09-770-315-3
16	109.4	5.4	8151	4	US-09-438-268-2
17	109.4	5.4	8179	4	US-09-438-268-5
18	109.4	5.4	8698	4	US-09-770-315-2
19	104.6	5.2	939	4	US-09-532-594B-12
20	104.6	5.2	1197	4	US-09-532-594B-13
21	104.6	5.2	1611	4	US-09-532-594B-14
22	104.6	5.2	1872	4	US-09-532-594B-3
23	104.6	5.2	1872	4	US-09-532-594B-15
24	104.6	5.2	4767	4	US-09-532-594B-1
25	99.8	4.9	969	4	US-09-807-802A-10
26	99.8	4.9	1200	4	US-09-807-802A-8
27	99.8	4.9	1641	4	US-09-807-802A-6

28	99.8	4.9	1872	4	US-09-807-802A-4	Sequence 4, Appli
29	99.8	4.9	4683	4	US-09-807-802A-19	Sequence 19, Appli
30	99.8	4.9	4718	4	US-09-807-802A-1	Sequence 1, Appli
C 31	64.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appli
C 32	61.2	3.0	1141	4	US-09-806-708B-22	Sequence 22, Appli
33	44	2.2	1141	4	US-09-806-708B-22	Sequence 22, Appli
34	40.2	2.0	277	3	US-09-007-005-3	Sequence 3, Appli
35	40.2	2.0	277	3	US-09-244-796-3	Sequence 3, Appli
C 36	39.2	1.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 37	39.2	1.9	1664976	4	US-09-692-570-1	Sequence 1, Appli
38	38.6	1.9	116592	4	US-09-818-512-3	Sequence 3, Appli
39	38.4	1.9	1367	4	US-09-370-767-14265	Sequence 14265, A
40	38.4	1.9	6328	3	US-08-913-832A-1	Sequence 1, Appli
41	38.4	1.9	6328	4	US-09-249-181A-1	Sequence 1, Appli
42	38.4	1.9	6328	4	US-09-158-707-1	Sequence 1, Appli
43	38.4	1.9	6475	4	US-09-620-312D-325	Sequence 325, App
44	37.8	1.9	399	4	US-09-621-976-8976	Sequence 8976, Ap
45	37.8	1.9	832	4	US-09-621-976-2813	Sequence-2813, Ap

#### ALIGNMENTS

RESULT 1  
US-08-336-345-1  
; Sequence 1, Application US/08336345  
; Patent No. 5814510  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,345  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30753  
; REFERENCE/DOCKET NUMBER: 7937-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Parvovirus  
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 5.5e-291;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;  
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTAAAGGAAAA 60

273 ATGCTCGCCACCGAGTATATCTGAGGAGTAAATGGTTAAAGAACAT 332  
61 AGTAACGAGGAGTGTCTCATTTGTTTTTAAATGCTCAACTGAATGAAAA 120  
333 GCAGAAATGAAGCATTTTGGTTGTTTTTAAATGTGACACGTCACATAATGAAG 392  
121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180  
393 GATGTTGCGTGGAAACAATATACCAACCAATTCAAAAATGAAGAGCTAACATCTTTAAAT 452  
181 CGAGGAGCGGAATCTACTTGGGACCAAGC-----GAGGACATGAATGGGAACCA 234  
453 AGAGGAGCACAACAGCAATGATCAACCCGAAAGAGAAATGAGCTGGGAATCGGA 512  
235 GTCGATGAATGACCAAAAGCAAGTATTCATTTTGGTTGTTTGAATTTTAAATGTTTA 294  
513 GTTGTATGCTCCGCAAAAGCAAGTACAACTTTTGTATGATTAATTAATAAATGTCTT 572  
295 TTTGAAAGTGTTAACCAAGAAATATATTTCTGGTGTATTTAAATGGTTTGTGCAACAT 354  
573 TTTGAAAGTCTTTCTTAAATATAGAACCAATGAATGTGTTGGTTTATTCACAT 632  
355 GAATGGGAAAAGACCAAGCTGGCACTGCCATGTACTAATTCGAGGAAGGACTTTAGT 414  
633 GAATGGGAAAAGATCAAGGCTGGCAATGTCTATGTTTACTTCATAGTAAGAACTTACAA 692  
415 CAAGCTCAAGGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGGTA 474  
693 CAAGCAACTGGTAATGGCTACGAGACAAATGAATATGTATTTGGAGTAGATGGTTGGTG 752  
475 ACAGCCTGTAATGTCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534  
753 ACTCTTTGTTGCGTAAACTTTAACCACTGAAAGATTAAGCTCAGAGAAATTCAGAA 812  
535 GACATAGAGTGGTTACTTACTTACTTATAGCATAGCAAAACCAAAAGAGACTATACC 594  
813 GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAAAACCAAAAGAGACTATGTT 872  
595 AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTTTAACTTAAAGAAATAAGC 654  
873 AATATGGTTCAITTTGGAAATATGATAGCATATTACTTTTAAAGAAATAATTTGTC 932  
655 ACTAGTCCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAATACTAAC 714  
933 CACAT--GACAAAGAAAGTGGCTATTTTAAAGTACTGATTCTGGTGGAAATTTAAC 989  
715 TTTTAAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCA 774  
990 TTTATGAAGTATCAAGACAGACAAATGTGCAGCACACTTTTACACTGAACAAATGAACCA 1049  
775 GAAACGGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 834  
1050 GAAACCGTTGAAACCAAGTACAGACAGACAGCAAAAGCGCGGAGAAATTCAACT 1109  
835 AAAAAAGAGTGTGCTATTAATACTTACATTTAAAGAGCTGGTGCATAAAAAGAGTAACTCA 894  
1110 AAAAAAGAGTGTCAATCAATGATGTTTTCGGGAGCTTGGTTAGTAAAGAGTAACTCA 1169  
895 CCAGAGGACTGGATGATGATGCGAGCCAGACAGTGTACATTTGAATGATGGCTCAACAGGT 954  
1170 CCTGAAGACTGGATGATGTTTACAAACCATAGATAGTTATATTGAATGATGGCAGCAACAGGA 1229  
955 GGAGAAACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTTAGCCAGAACCAAA 1014  
1230 GGTGAATCTTTTAAAAATATCACTTGAATTTGTACTTTGACTTTAGCAGAACCAAAA 1289  
1015 ACAGCATTTGACTTAATTTTAAAGAAAGCTGAAACCAAGCAAACTAACACTTTTCACTG 1074  
1290 ACAGCATTTGAAATTAATCTTGAAGAGCAGATAACTAACTAACTAACTTTGATCTT 1349  
1075 CTTGACACAGAACTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTGC 1134

1350 GCAAATTTAGAACATGTCAAATTTTAAATGTCACGGATGGAATTTGGATTTAAAGTTTGT 1409  
1135 CATGCTATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTTATTTTCAT 1194  
1410 CACGCTATAGCATGTGTTTTTAAATAGACAAGGTGGTAAAGAAATAACAGTCTCTTTTTCAT 1469  
1195 GACACGCGCAGACAGGCAAAATCTATTATTGACAAGCCATAGCACAAAGCAGTTGCAAT 1254  
1470 GGACCAAGCAAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCACAAAGCTGTGGTAAT 1529  
1255 GTTGGTTGCTATAATGTCAGCCAAATGTAAATCTTTCAATTTAAATGACTGTACCAACAAGAAC 1314  
1530 GTTGGTTGTTAATGTCAGCAAAATGTAAATTTTCCATTTAATGACTGTACCAATAAAT 1589  
1315 TTGATTTGGGTAGAGAAGCTGTAACTTTTGACAGCAAGTAACCAAGTTTAAAGCCATT 1374  
1590 TTAATTTGGATTGAAGAAGCTGTAACTTTTGGTCAACAAGTTAACTCAATTTTAAAGCAATC 1649  
1375 TGTCTGTGTCAAACTATTTCGCAATTTGATCAAAAAGGAAAGGCAAGCAACAGATTGAACCA 1434  
1650 TGTCTGAGCAAACTATTAGAAATTTGATCAAAAAGGTAAGAGGAAATTTGAACCA 1709  
1435 ACACCAAGTCATCATGACCAACAATGAGAACATTTACAGTGTGTCAGAAATAGCTGCGAAGAA 1494  
1710 ACTCCAGTAAATTTAGCAACTAATGAAATATAACAATTTGTGGAATTTGATGTGAAGAA 1769  
1495 AGACCAAGAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAACACATACC 1554  
1770 AGACCTGAACATACACAACCAATTAAGAGACAGAAATGTTGAAACATTAAGTTAGTATGAAG 1829  
1555 TTGCTGTGTGACTTTGGTTTGGTTGACAAAAATGAATGCCCCCATGATTTGTGCTGGTTG 1614  
1830 CTTCCAGGAGACTTTGGTTTGGTTGATAAAGAGAAATGCGCTTTAATATGTGCATGGTTA 1889  
1615 GTAAAGAAATGGTTTACCACATCTACCATGCGCAAGCTACTGTGCTAAATGGGCAAAATTCCT 1674  
1890 GTTAAACATGGTTTGAATCAACCATGGCTACTATACATCATTTGGGGAAGAAAGTACCA 1949  
1675 GATGCTCAGAAAACTGGCGGAGGCAAAAGGTGCAAACTCTTATATAAATTTACTAGTTGCG 1734  
1950 GAATGGATGAAAACTGGCGGAGCCTAAAAATACAGAAAGGTATAAATTTCAACGAGTTGC 2009  
1735 GACGCTCACCATTACAGACACGAAAGAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794  
2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGCAAGTTCTAACT 2066  
1795 CCACTTGCATCGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAAAATACTCCT 1854  
2067 CCTCTGACTCCGAGCGTAGTGGACCTTGCACTGGAAACCGTGGAGTACTCCAGATACGCT 2126  
1855 GTTGGCGGCACTGCAGAAAAACCCAGAACACTGGGGAGAGCTGGTTTCCAAAGCTGCCAAGAT 1914  
2127 ATTGCAAGAACTGCAAAATCAACAACTCAAACTTTGGGCTTACTCA-----CAAAAGAC 2180  
1915 GGTCAACTGAGCCCACTTTGGTCAGAGATCGAGGAGATTTGAGAGGCTGCTTCGGTGG 1974  
2181 GTGCAAGCAGAGTCCGAGCTGGTCCGAAATAGAGGACAGCTGAGAGCAATCTTACTTCT 2240  
1975 GAACCTTTGAAGAAAGACTTTTCAGCGA 2000  
2241 GAACAATTTGAAGAAATTTTCGAGA 2266

## RESULT 2

US-08-336-345-2

; Sequence 2, Application US/08336345

; Patent No. 5814510

; GENERAL INFORMATION:

; APPLICANT: Parrish, Colin R.

; APPLICANT: Gruenberg, Allen

; APPLICANT: Carmichael, Leland E.

; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine

; NUMBER OF SEQUENCES: 2





1615 GTAAAGATGGTTACCAATCTTACCATGGCAAGCTACTGTCTAAATGGGCAAAAGTTCCCT 1674  
1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACACATCATTTGGGGAAGTACCA 1949  
1675 GATTGGTCAGAAACTGGGGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGGTTGC 1734  
1950 GAATGGGATGAAACTGGGGGAGCCCTAAAATACAAAGAGGTATAAATTCACCGAGTTGC 2009  
1735 GCAGCTCACCATTTCAGCACACCGAAGTACCGCTCTCAGCCAGACTATGCACTAACT 1794  
2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066  
1795 CCACTTCGATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATATCTCT 1854  
2067 CCTCTGACTCGGAGCTAGTGGACCTTGCATCGAACCGTGGAGTACTCCAGATACGCT 2126  
1855 GTTGGCGGCACTGAGAAACCCAGAACACTCTGGGGAAGCTGGTTCCTCAAGCCTCCAGAT 1914  
2127 ATTGCAGAACTGCAAAATCAAAATCAAACTCAAACTTGGCGTTACTCA-----CAAAGAC 2180  
1915 GGTCAACTGAGCCCAACTCTGTCAGAGATCGAGGAGATTGGAGCGTCTTCCGTTGG 1974  
2181 GTGCAACGAGTCCGAGTGGTCCGAATAGAGCGACCTGAGAGCCATCTTTACTTCT 2240  
1975 GAACCGTTGAAGAAAGACTTTCAGCGA 2000  
2241 GAACAATTGGAAGAGATTTTCGAGA 2266

RESULT 3  
US-08-647-655-1  
; Sequence 1, Application US/08647655  
; Patent No. 5885585  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/647,655  
; FILING DATE: On Even Date Herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30,753  
; REFERENCE/DOCKET NUMBER: 7937-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOHETICAL: NO  
; ANTI-SENSE: NO  
; ORGANISM: Parvovirus

US-08-647-655-1

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 5.5e-291;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60  
DB 273 ATGTCTGGCAACCCAGTATATCTAGGAAGTTATGGAGGAGTAAATTTGGTTAAAGAAACAT 332  
QY 61 AGTAACCGAAGAGTGTCTCATTTGTTTTAAAAATGAAAATGTTCACTGAATGGAATA 120  
DB 333 GCAGAAATGAAAGCAATTTTCGTTGTTTTAAATGTGCAACGTCCTCACTAAATGGAAG 392  
QY 121 GATATCGGATGAAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTACAA 180  
DB 393 GATGTTGCTGGTGAACCACTATACCAAA CCAATTCAAATGAAAGAGCTACATCTTTTAAT 452  
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAAAGC-----GAGGACATGGAATGGGAAACCA 234  
DB 453 AGAGGAGCACAAACAGCAATGATCAACCCGAAAGAAAGAAATGAGACTGGGAATCGGAA 512  
QY 235 GTGATGAAATGACCAAAAGCAAGTATTCATTTTGTGTTTGTGTTTAAATAATGTTTA 294  
DB 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGTGATGCAATTAATTTAAAAAATGTCT 572  
QY 295 TTTGAAGTGTCTTAAACACAAAGAAATATATTTCTGCTGATGTTTAAATTTGTTGTAACAT 354  
DB 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAATGATGTTGTTGTTTATTCACAT 632  
QY 355 GAATGGGAAAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGACTTTTACT 414  
DB 633 GAATGGGAAAAAGATCAAGCTGGCATTGTCTGTTTACTTTCATAGTAAGAACTTACAA 692  
QY 415 CAACTCAAGGAAATGTGGAAGAGGCACTAAATGTTTACTGAGCAGATGTTGGTA 474  
DB 693 CAAAGCACTGGTAAATGGCTACGACACAAATGAAATATGTATTTGGAGTAGATGGTTGGTG 752  
QY 475 ACAGCTGTAATGTGCACTAACACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAA 534  
DB 753 ACTCTTTGTTGCGTAAACTTAACCACTGAAAGATTTAGCTCAGAGAAATTCAGAA 812  
QY 535 GCAATGAGTGGTTACTCTACTTATTATAGCATAAGCAAAACCAAAAGACTATACC 594  
DB 813 GATAGTGAATGGTGTACTATATTAACATACAGACATAAGCAAAACCAAAAGACTATGTT 872  
QY 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATTAAGC 654  
DB 873 AAAATGGTTCAATTTTGGAAATATGATAGCATATTACTTTTTTAAAGAAATTTGTC 932  
QY 655 ACTAGTCCACCAAGAGACGAGGCTTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAAC 714  
DB 933 CACAT---GACAAAGAAAGTGGCTATTTTTTAAGTACTGATTTCTGTTGGAATTTAAC 989  
QY 715 TTTTAAAGAAAGGCGAGCGCCATCTAGTGACAACTATACATGATGACATCGGGCCA 774  
DB 990 TTTATGAAGTATCAAGACAGACAAATTTGTACACACACTTTTACACTGAACAAATGAACCA 1049  
QY 775 GAAACGGTTGAACCAACAGTAACTGCGGAGGAACTAAGCGGCGGAGATTTCAACT 834  
DB 1050 GAAACCGTTGAACCAACAGTAACTGCGGAGGAACTAAGCGGCGGAGATTTCAACT 1109  
QY 835 AAAAAAGAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCA 894  
DB 1110 AAAAGGAAGTGTCAATCAATGTAATTTTGGGACTTTGGTTAGTAAAGAGTAACTCA 1169  
QY 895 CCAGAGGACTGGATGATGATGACGACAGTGTACATTTGAATGATGGCTCAACAGGT 954  
DB 1170 CCTGAAGACTGGATGATGTTACCAACAGATAGTTATATTGAAATGATGCGCAACACGAG 1229  
QY 955 GGAGAAACCTGCTGAAAAATACCGTAGAGATTTGTACTACTTCTAGCCAGAACCCAA 1014  
DB 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTTGTACTTTGACTTTTACAGAAACAA 1289

Qy	1015	ACGCACTTTGCATTAAATTTTTAGAAAAAGCTGAACACGACGAAAATAACCAACTTTTCACTCG	1074
Dd	1290	ACGCACTTTGAAATTAATCTTTGAAAAAGCAGATAAATACTAACTAACTAACTTTGATCTTT	1349
Qy	1075	CCTTGACAACAAGAACCTCGAGAAATTTTTTGGCTTTTCATGGCTGGAACTATGTATAAGTTTGC	1134
Dd	1350	GCAAATCTTAGAACATGCTCAAAATTTTTTAGAATGCAACGGATGGAATTCGATTAAGTTTGT	1409
Qy	1135	CATGCTAATTTGCTGTGTTTTTAACACAGACAGGAGGCCAAAANAATACTGTTTATTTCAT	1194
Dd	1410	CAGCTATAGCATGTGTTTTTAAATAGACAAGGTGGTAAAGAANAATACAGTTCCTTTTTCAT	1469
Qy	1195	GGACCAGCCAGCACAGGCAAACTCTATTATTATGCAACAAGCCATAGCACAAAGCAGTTGGCAAT	1254
Dd	1470	GGACCACAGTAGCAGGNAATCTATCATTTGCTCAAGCCATAGCACAAAGCTGTGGGTAAAT	1529
Qy	1255	GTTGGTTGCTATTAATGCAAGCAATGTAAACTTTTCCTATTTAATGACTGTACCAACAAGAAC	1314
Dd	1530	GTTGGTTGTTATTAATGCAAGCAAAATGTAATTTTTTCCATTTAATGACTGTACCAATAAAAT	1589
Qy	1315	TTGATTTGGGTAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATT	1374
Dd	1590	TTTAATTTGGATGGAAGAACTGGTAACTTTGGTCAACAGTTAATCAATTTAAAGCAATC	1649
Qy	1375	TGCTCTGGTCAAACTATTTCGCATTTGATCAAAAAAGGAAAAGCCAGCAAAACAGATTGAACCA	1434
Dd	1650	TGTTCTGGAACAAACAAATTAGAAATTTGATCAAAAAGGTAAAGAAAGTAGCAAAATTGAACCA	1709
Qy	1435	ACACCACTCATATGACCAACAAATGAGAAACATTTACAGTGTCTAGAAATAGCTCGGAGAA	1494
Dd	1710	ACTCCAGTAATTTATGACAACTAATGAAAAATATAACAAATTTGTAGAAATTTGATGTGAAGAA	1769
Qy	1495	AGACCAAGAACACACTCAACCAATCAGACAGACAGAAATGCTTAAACATTCATCTAAACACATACC	1554
Dd	1770	AGACCTGAAACATACACAAACCAATAAGACAGACAATGTTTGAACATTAAGTTAGTATGTAAG	1829
Qy	1555	TTGCTCGTGACATTTGGTTTGGTTGTGACAAAAATGAAATGGCCCATGATTTGTGCTTGGTTG	1614
Dd	1830	CTTCCAGGACATTTGGTTTGGTTGTATTAAGAAGAATGGCCTTTAAATATGTGCATGGTTA	1889
Qy	1615	GTAAGAATAGTTTACCAATCTACCATGGCAAGCTTACTGTGCTAAATGGGGCAAAAGTTCCCT	1674
Dd	1890	GTTTAAACATGGTTTGAATCAACCATGGCTTAACTATACACATCATTTGGGGAAGATACCA	1949
Qy	1675	GATTGGTCAGAAAACTGGGCGGAGCCAAAAGGTGCAACTCCTATAATTTACTAGGTTGCG	1734
Dd	1950	GAATGGGATGAAAACTGGGCGGAGCCTTAAATAACAAGAAGGTATAATTCACCAGGTTGC	2009
Qy	1735	GCAGCTCACCATTTACGACAGACACGGAAAGTAGCGCTCTCGACCGAGAATATGCACTAACT	1794
Dd	2010	AAGACT---TAGAGACACAAAGCGGCAAGCAANTCCTCAGAGTCAAGACCAAGTTCTTAACT	2066
Qy	1795	CCACTTGCACTCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACAAAATPACTCCT	1854
Dd	2067	CCTCTGACTCCGACGCTAGTGGACCTTTGCACTGGAAACCGTGGAGTACTCCAGATACGCT	2126
Qy	1855	GTTGCGGGCACTGCAGAAAACCCAGAAACACTGGGGGAAGCTGGTTCCAAGGCTTCCCRAGAT	1914
Dd	2127	ATTGCAAAAACCTGCAAAATCAACAAATCAAAACCACTTGGCGTTACTCA-----CAAAGAC	2180
Qy	1915	GGTCAACTGAGCCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGGTGCTTCGGTGGC	1974
Dd	2181	GTGCAAGCGAGTCGACGTTGGTCGGAATAGAGGCGACACCTGAGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTTGAAGAAAGACTTTCAGCGA	2000
Dd	2241	GAAACAATTGGAAGAGATTTTCGAGA	2266

```

: Patent No. 5885585
:
: GENERAL INFORMATION:
:
: APPLICANT: Parrish, Colin R.
: APPLICANT: Gruenberg, Allen
: APPLICANT: Carmichael, Leland E.
: TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
:
: NUMBER OF SEQUENCES: 2
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: ZIP: 10036-2711
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/647,655
: FILING DATE: On Even Date Herewith
:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Jennifer
: REGISTRATION NUMBER: 30,753
: REFERENCE/DOCKET NUMBER: 7937-008
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
:
: INFORMATION FOR SEQ ID NO: 2:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5049 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: unknown
:
: MOLECULE TYPE: DNA (genomic)
:
: HYPOTHETICAL: NO
:
: ANTI-SENSE: NO
:
: ORIGINAL SOURCE:
: ORGANISM: Parvovirus
:
: US-08-647-655-2

```

Query Match	51.5%;	Score 1040.4;	DB 2;	Length 5049;
Best Local Similarity	71.4%;	Pred. No. 5.5e-391;		
Matches 1492;	Conservative 0;	Mismatches 556;	Indels 18;	Gaps 4;
Qy	1	ATGCCTGGAAATGCTTACTCTGATGAAGTGTTCGGAGCAACCAACTGGTTAAAGGAAAAA	60	
Db	273	ATGCTGGCAACAGTATACTGAGGAAGTATGAGGGAGTAATTGGTTAAGAAACAT	332	
Qy	61	AGTAACACAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGGAAAA	120	
Db	333	GCAGAAAAATGAAGCAATTTTCGTTGTGTTTTTAAAAATGTCAACACGTCCTCAACTGAAAG	392	
Qy	121	GATATCGGATGGAAATAGTTTACAAAAGAAGAGCTGCAGGAGGACGAGCTGGAATCTTTTACAA	180	
Db	393	GATGTTCGCTGGAAACAACTATACCAACCAATTCAAAATGGAAGAGCTAACATCTTTAATT	452	
Qy	181	CGAGGAGCGGAAACTACTTGGGACCAAGC-----GAGGACATGGAAATCGGAAACCACA	234	
Db	453	AGAGGAGCAACAACAGCAATGGATCAAAACGGAAGAGAGAAATGGACTGGGAATCGAA	512	
Qy	235	GTGGATGAAATGACCAAAAAGCAAGTATTCATTTTTTGTATTCCTTTGGTTAAAAAATGTTTA	294	
Db	513	GTTTGATAGTCTGCCCAAAAAGCAAGTACAAACTTTTTGTATGCATTTAATTTAAAAAATGTCCT	572	
Qy	295	TTTGAAGTGTCTTAACACAAAGAATATATTTCTCTGGTGTATTTAATTTGGTTTGTGCAACAT	354	
Db	573	TTTGAAGTCTTTGTTCTAAAAATATAGAACCAATGAATGTGTTTGGTTTATTCACAT	632	
Qy	355	GAATGGGAAAAAGACCAAGGCTGGCACTGCCATGTACTAATATGGAGAAAGGACTTTAGT	414	
Db	633	GAATGGGAAAAAGATCAAGCTGGCAATGTCATGTTTTTACTTCATAGTAAAGAACTTACAA	692	

RESULT 4  
US-08-647-655-2  
; Sequence 2, Application US/08647655

QY	415	CAAGCTCAAGGGAAATGGTGGAGAAAGGCAACTAAATGTTTACTCTGGACGAGATGGTTCGGTA	474
DB	693	CAAGCAACTGGTAAATGGCTACGACAGACAAATGAATATGTAATTCGGACTAGATGGTTCGGTG	752
QY	475	ACAGCCGTGAATGTGCACAACTAAACACCAAGCTGAAAGAAATTAACCTAAACAGAGAAATAGCAGAA	534
DB	753	ACTCTTTGTCGGTAAACTTAACACCAACTGAAAGAATTAAGCTCAGAGAAATTCGAGAA	812
QY	535	GACAATGAGTGGGTACTCTACTTACTTATAAGCATAAAGCAAAACCAAAAAAGACTATATACC	594
DB	813	GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAACAAACAAAGACTATGTT	872
QY	595	AAGTGTGTTCTTTTGGAAACATGATTCCTTAATAATTTTTTAACATAAAAAAGAAATAAGC	654
DB	873	AAAATGTTTCATTTTGGAAATATGATAGCATATTACTTTTTTAAACAAAGAAAAAAATGTC	932
QY	655	ACTAGTCCACACAGAGACGGGCTATTTCTTTAGCAGTGACTCTGGCTCGAAAACTAAC	714
DB	933	CACAT---GACAAAGAAAGTGGCTATTTTTTAAGTACTGATCTGGTTCGGAATTTTAAC	989
QY	715	TTTTTAAAGAGCGCGCCACTCTAGTGAGCAAACTATACATGATGACATCGCGGCCA	774
DB	990	TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGAACCA	1049
QY	775	GAAACGGTTGAACACCAAGTAAACACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACT	834
DB	1050	GAAACCGTTGAAACCAAGTGACACAGCACAGGAAACAAAGCGCGGGGAATTCAAACT	1109
QY	835	AAAAAGAAAGTGTGCTATTTAAAACTTACACTTTAAAGAGCTGCTGCATAAACAGATAACCTCA	894
DB	1110	AAAAAGAAAGTGTCAATCAAAATGTACTTTGCGGGACTTTAGTTAGTAAAGATTAACATCA	1169
QY	895	CCAGAGACTGGATGATGATGCAGCCAGACAGTTTACATTGAAATGATGGCTCAACCAGGT	954
DB	1170	CCTGAAGACTGGATGATGTTTACAACACAGATAGTTATATTGAAATGATGGCACAAACGAGGA	1229
QY	955	GGAGAAACCTGCTGAAAAATACGCTAGAGATTTGTACATACTCTAGCCAGAACCAAA	1014
DB	1230	GGTGAAATCTTTTAAAAAATACACTTGAATTTGTACTTTTGAATTTAGCAAGAACAAAA	1289
QY	1015	ACAGCATTTGCATTAAATTTTAGAAAAAGCTGAAACACAGCAAACTAACCAACTTTTCACTG	1074
DB	1290	ACAGCATTTGAAATTAATCTTTGAAAAGCAGATAATACTAACTAACTACTTTTGATCTTT	1349
QY	1075	CCTGACACAAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTGC	1134
DB	1350	GCAAAATCTAGAACATGTCAAAATTTTAGAAATGACGAGTGGAAATTTGATTAAGTTGT	1409
QY	1135	CATGCTATTTGCTGTGTTTTTAAACAGACAGGAGGCAAGAAATACTGTTTTTATTTTCAT	1194
DB	1410	CAGCCTATAGCATGTGTTTTTAAATAGACAAGTGGTAAAAAGAAATACAGTTCTTTTTCAT	1469
QY	1195	GGACCAAGCAGCACAGCAAACTCTATTATTGCAACAGCCATAGCACAAAGCAGTTGGCAAT	1254
DB	1470	GGACCAAGCAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCACAAAGCTGTGGTAAAT	1529
QY	1255	GTTGGTTGCTATAATGCAGCAATGTAAACTTTCCATTTTAATGACTGTACCAACAAGAAC	1314
DB	1530	GTTGGTTGTTAATGCAGCAATGTAAATTTTCCATTTTAATGACTGTACCAATAAAAAAT	1589
QY	1315	TTGATTTTGGGTAGNAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAGCCATT	1374
DB	1590	TTAAATTTGGAATGAAGAAGCTGTAACTTTGGTCAACAGTTAATCAATTTAAAGCAATC	1649
QY	1375	TGCTCTGGTCAAACTATTTCCATTTGATCAAAAAAGGAGGACGACAAAACAGTTGAACCA	1434
DB	1650	TGTTCTGGACAAACAATTAGAAATGATCAAAAGGTAAGGAAGTAGCAAAATTTGAACCA	1709
QY	1435	ACAACGATCATATGACCACAAAATGAGAACATTTACGTGGTCAAGATAGGCTCGGAAGAA	1494
DB	1710	ACTCCAGTAATTATGACCAATAATGAAAAATATACAAATTTGAGAAATTTGATGTGAAGAA	1769

## RESULT 5

US-08-254-358-1

US 00-254-338-1  
; Sequence 1, Application US/08254358

; Patent No. 5658785

; GENERAL INFORMATION:

APPLICANT: Johnson, Philip R.

**TITLE OF INVENTION:** Adeno-Associated Virus Materials and

;	TITLE OF INVENTION:	AUTOMATIC
:	TITLE OF INVENTION:	METHODS

: TITLE OF INVENTION: MECHANICAL  
 : NUMBER OF SEQUENCES: 3

; NUMBER OF SEQUENCES: 3  
: CORRESPONDENCE ADDRESS:

;; CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray  
STREET: 6300 Sears Tower 233 S Wacker Drive

STREET: 6300  
CITY: Chicago

CLTY: Chicago  
STATE: Illinois

STATE: Illinois  
COUNTRY: USA

; COUNTRY: U  
ZIP: CACAC

ZIP: 60606

; COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: Floppy disk
;

```

```

;
;
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Rel

; CURRENT APPLICATION DATA:

APPLICATION

; FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5658785and, Gr

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31

TELECOMMUNICATION INFORMATION

TELEPHONE: (312) 474-6311

TELEPHONE: (312) 312-1111  
TELEFAX: (312) 312-1111

TELEX: 25-3856

TELEX. 23-3030  
: INFORMATION FOR SEO ID NO:

SEQUENCE CHARACTERISTICS:

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match      5.4%; Score 109.4; DB 1; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGGAGGAAATCTGTTTATTTTCATGGACCGACGACGACGCAATCTATTATTGCA 1227
Db      |||||
QY 1293 GGCAGGAGGAAATCTGTTTATTTTCATGGACCGACGACGACGCAATCTATTATTGCA 1352
Db      |||||
QY 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTGTCTATTAATGCAGCCAAATGTAACATTT 1287
Db      |||||
QY 1353 GAGGCCATAGCCACACTGTGCTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1412
Db      |||||
QY 1288 CCATTTAATCACTGTACCAACAGAACTTGAATTTGGTGTCTATTAATGCAGCCAAATGTAACATTT 1347
Db      |||||
QY 1413 CCCTTCAACGACTGTGCGAAGATGGTGTATCTGGTGGAGGAGGGAAGATGACCGCC 1472
Db      |||||
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCAATTTGATCAAAA 1407
Db      |||||
QY 1473 AAGTCTGTGAGTGGCGCAAGCCATTTCTCGGAGGAGCAGGTGCGGTGGACCAAAA 1532
Db      |||||
QY 1408 GGAAGGAGGAGCAACAGATTGAACCAACACCATCATCATGACCAAAATGAGAACATTT 1467
Db      |||||
QY 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACTCCAAACCAACATG 1592
Db      |||||
QY 1468 ACAGTGTGAGATAGCTCGGAGGAAAGACCAAGACCACTCAACCAATCAGACAGAGA 1527
Db      |||||
QY 1593 TGCCTGCTGATTTGACGGAACTCAACGACCTTCGACACACAGCAGCGGTTGCAAGACCGG 1652
Db      |||||
QY 1528 ATGCTTAAACATTCATCTAAACATACCTTCCTGCTGACTTTGTTGGTTGACAAAAT 1587
Db      |||||
QY 1653 ATGTTCAAAATTTGAATCTACCCCGCTCTGATCATGACTTTGGGAGGTTGACCAAGCAG 1712
Db      |||||
```

## RESULT 6

```
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.
```

```
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-475-391-1
```

```
Query Match      5.4%; Score 109.4; DB 1; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGGAGGAAATCTGTTTATTTTCATGGACCGACGACGACGCAATCTATTATTGCA 1227
Db      |||||
QY 1293 GGCAGGAGGAAATCTGTTTATTTTCATGGACCGACGACGACGCAATCTATTATTGCA 1352
Db      |||||
QY 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTGTCTATTAATGCAGCCAAATGTAACATTT 1287
Db      |||||
QY 1353 GAGGCCATAGCCACACTGTGCTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1412
Db      |||||
QY 1288 CCATTTAATCACTGTACCAACAGAACTTGAATTTGGTGTCTATTAATGCAGCCAAATGTAACATTT 1347
Db      |||||
QY 1413 CCCTTCAACGACTGTGCGAAGATGGTGTATCTGGTGGAGGAGGGAAGATGACCGCC 1472
Db      |||||
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCAATTTGATCAAAA 1407
Db      |||||
QY 1473 AAGTCTGTGAGTGGCGCAAGCCATTTCTCGGAGGAGCAGGTGCGGTGGACCAAAA 1532
Db      |||||
QY 1408 GGAAGGAGGAGCAACAGATTGAACCAACACCATCATCATGACCAAAATGAGAACATTT 1467
Db      |||||
QY 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACTCCAAACCAACATG 1592
Db      |||||
QY 1468 ACAGTGTGAGATAGCTCGGAGGAAAGACCAAGACCACTCAACCAATCAGACAGAGA 1527
Db      |||||
QY 1593 TGCCTGCTGATTTGACGGAACTCAACGACCTTCGACACACAGCAGCGGTTGCAAGACCGG 1652
Db      |||||
QY 1528 ATGCTTAAACATTCATCTAAACATACCTTCCTGCTGACTTTGTTGGTTGACAAAAT 1587
Db      |||||
QY 1653 ATGTTCAAAATTTGAATCTACCCCGCTCTGATCATGACTTTGGGAGGTTGACCAAGCAG 1712
Db      |||||
```

## RESULT 7

```
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
```

Mon Jan 24 09:41:23 2005

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,609
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 585877sand, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-709-609-1

Query Match 5.4%; Score 109.4; DB 2; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGAAAGAAATACGTGTTTATTTTCATGGACCCAGCCAGCAGCAAGCAATCTATTATTGCA 1227
DB 1293 GGCAGAGGAAACCATCTGGCTGTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCACAAAGAGTTGGCAATGTTGGTTGCTATATATGACGCCCAATGTAACCTTT 1287
DB 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGGCTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAAGACTTGAATTTGGGTAGAGAAAGCTGTTACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTCGAAAGATGGTGATCTGGTGGAGAGGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACCAAGTTAAAGCCATTGCTCTGGTCAAACTATTGGCATTTGATCAAAA 1407
DB 1473 AAGTCTGGAGTCGGCAAGCCATTCTCGGAGGAAGCAAGTGCGCGTGGACAGAA 1532
QY 1408 GGAAGGAGCAAAACAGATGTAACCAACCAAGTGAAGCAACACACTCAACCAATCAGAGACA 1527
DB 1593 TGCGCCGTGATTGACGGGAATCAACGACCTTCGAACACCAAGAGCGCTTGGAGAGGTCAACAGCAG 1652
QY 1528 ATGCTTAACATTATCTAACACATACCTTGCTGGTGAATTTGGTTTGGTGAACAAAAT 1587
DB 1653 ATGTTCAAAATTGAATCACCAGCGCTGATGATCATGACTTTGGAGAGGTCAACAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

```
RESULT 8
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and Methods
; NUMBER OF INVENTION: 3
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07178
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US95-07178-1

Query Match 5.4%; Score 109.4; DB 5; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGAAAGAAATACGTGTTTATTTTCATGGACCCAGCCAGCAGCAAGCAATCTATTATTGCA 1227
DB 1293 GGCAGAGGAAACCATCTGGCTGTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCACAAAGCTTGGCAATGTTGGTTGCTATATATGACGCCCAATGTAACCTTT 1287
DB 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGGTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAAGACTTGAATTTGGGTAGAGAAAGCTGTTACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTCGAAAGATGGTGATCTGGTGGAGAGGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACCAAGTTAAAGCCATTGCTCTGGTCAAACTATTGGCATTTGATCAAAA 1407
DB 1473 AAGTCTGGAGTCGGCAAGCCATTCTCGGAGGAAGCAAGTGCGCGTGGACAGAA 1532
QY 1408 GGAAGGAGCAAAACAGATGTAACCAACCAAGTGAAGCAACACACTCAACCAATCAGAGACA 1467
DB 1593 TGCGCCGTGATTGACGGGAATCAACGACCTTCGAACACCAAGAGCGCTTGGAGAGGTCAACAGCAG 1652
QY 1528 ATGCTTAACATTATCTAACACATACCTTGCTGGTGAATTTGGTTTGGTGAACAAAAT 1587
DB 1653 ATGTTCAAAATTGAATCACCAGCGCTGATGATCATGACTTTGGAGAGGTCAACAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

```
RESULT 9
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences, Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVNP.031USA
```

; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 18  
; TYPE: DNA  
; ORGANISM: AAV-2  
US-09-807-802A-18

Query Match 5.4%; Score 109.4; DB 4; Length 4681;  
Best Local Similarity 53.7%; Pred. No. 4e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGAGGAAGTAATCTGTTTATTTTCATGGACCCAGCCAGCACAGGCAAAATCTATTATTGCA 1227  
DB 1293 GGCAGAGGAAGTAATCTGTTTATTTTCATGGACCCAGCCAGCACAGGCAAAATCTATTATTGCA 1352  
QY 1228 CAAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCAAATGTAATCTTT 1287  
DB 1353 GAGCCATAGCCACACACTGTCGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTTT 1412  
QY 1288 CCATTAAATGACTGTACCAACAAGAACTTGATTGGGTAGAAAGAGCTGTAATCTTTGGA 1347  
DB 1413 CCCTTCAACGACTGTGTCGCAAGATGTTGATCTGGTGGGAGGAGGAGATGACCGCC 1472  
QY 1348 CAGCAAGTAAACGAGTTAAAGCCATTTGCTCTGGTCAAACTATTGCAATGATCAAAAA 1407  
DB 1473 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGGCGGTGGACCAAGAA 1532  
QY 1408 GGAAGAGGAGCAAAACAGATTGACCAACCAAGTATCATATGACCAACAATGAGAACATT 1467  
DB 1533 TGCAGTCTTCGGCCAGATAGACCGACTCCCGTATGTCACCTCCACACCAACATG 1592  
QY 1468 ACAGTGGTCAGATAGCTCGGAAGAAAGACAGAACACACTCAACCAATCAGAGACAGA 1527  
DB 1593 TGGCCGCTGATTACGCGGAATCAACGACCTTCGAACACACAGCAGCGTTGCAAGACCG 1652  
QY 1528 ATGCTTAAACATTCATCAACACATACCTTGCCTGGTACCTTTGGTTGGTGAACAAAT 1587  
DB 1653 ATGTTCAAATTTGAATCACTCCCGCCGCTCTGGATCATGACTTTTGGGAAGGTCAACCAAGCAG 1712  
QY 1588 GAA 1590  
DB 1713 GAA 1715

RESULT 10  
US-08-331-384-2/c  
; Sequence 2, Application US/08331384  
; Patent No. 5856152  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Kelley, William M.  
; APPLICANT: Fisher, Krishna J.  
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods  
; TITLE OF INVENTION: of Use Thereof  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/331,384  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: UPNG1149USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4910 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
US-08-331-384-2

Query Match 5.4%; Score 109.4; DB 2; Length 4910;  
Best Local Similarity 53.7%; Pred. No. 4.2e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGAGGAAGTAATCTGTTTATTTTCATGGACCCAGCCAGCACAGGCAAAATCTATTATTGCA 1227  
DB 1544 GGCAGAGGAAGTAATCTGTTTATTTTCATGGACCCAGCCAGCACAGGCAAAATCTATTATTGCA 1485  
QY 1228 CAAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCAAATGTAATCTTT 1287  
DB 1484 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1425  
QY 1288 CCATTAAATGACTGTACCAACAAGAACTTGATTGGGTAGAAAGCTGTAATCTTTGGA 1347  
DB 1424 CCCTTCAACGACTGTGTCGCAAGATGTTGATCTGGTGGGAGGAGGAGATGACCGCC 1365  
QY 1348 CAGCAAGTAAACGAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCAATGATCAAAAA 1407  
DB 1364 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGGCGGTGGACCAAGAA 1305  
QY 1408 GGAAGAGGAGCAAAACAGATTGAAACCAACACAGTATCATGACCAACAATGAGAACATT 1467  
DB 1304 TGCAGTCTTCGGCCAGATAGACCGACTCCCGTATGTCACCTCCACACCAACATG 1245  
QY 1468 ACAGTGGTCAGATAGCTCGGAAGAAAGACAGAACACACTCAACCAATCAGAGACAGA 1527  
DB 1244 TGGCCGCTGATTGACGGGAATCAACGACCTTCGAACACAGCAGCGTTGCAAGACCG 1185  
QY 1528 ATGCTTAAACATTCATCAACACATACCTTGCCTGGTACCTTTGGTTGGTGAACAAAT 1587  
DB 1184 ATGTTCAAATTTGAATCACTCCCGCCGCTCTGGATCATGACTTTTGGGAAGGTCAACCAAGCAG 1125  
QY 1588 GAA 1590  
DB 1124 GAA 1122

RESULT 11  
US-08-836-087-2/c  
; Sequence 2, Application US/08836087  
; Patent No. 5871982  
; GENERAL INFORMATION:  
; APPLICANT: Trustees of University of Pennsylvania  
; APPLICANT: Wilson, James M.  
; APPLICANT: Kelley, William M.  
; APPLICANT: Fisher, Krishna J.  
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
; TITLE OF INVENTION: Methods of Use Thereof  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, PO Box 457



```

; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,087
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/331,384
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVFN.007PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; US-08-836-087-2

Query Match          5.4%; Score 109.4; DB 2; Length 4910;
Best Local Similarity 53.7%; Pred No. 4,2e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 1168 GCGAAAGAAATACTGTTTTATTTTCATGGACCGCCAGCAGACGACAAATCTATTATTGCA 1227
Db      |||||
1544 GGCAAGAGGAACACCAATCGCTGTTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1485
Qy 1228 CAGCCATAGCAGACGAGCTTGGCAATGTTGGTTGCTATTAATGCAGCCAATGTAAACTTT 1287
Db      |||||
1484 GAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1425
Qy 1288 CCATTTAATGACTGTACCAACAACTTGATTTGGGTAGAGAAGCTGGTAACCTTTGGA 1347
Db      |||||
1424 CCCTTCAAGACTGTGTGCAAGATGGTGATCTGGTGGAGGAGGGAAGATGACGCC 1365
Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTGCTTGGTCAAACTATTCGATTTGATCAAAA 1407
Db      |||||
1364 AAGTCTGGAGTCGGCCAAAGCCATCTCGGAGGAAGCAAGGTGCGGTGGACCAAAA 1305
Qy 1408 GGAAAGGCGAGCAACAGATTGAACCAACACCACTCATATGACCACAAATGAGAACATT 1467
Db      |||||
1304 TGCAAGTCCTCGGGCCAGATAGACCCGACTCCCGTGATCGTCACTCCAACCAACATG 1245
Qy 1468 ACAGTGTCTCAATAGGCTCGAAGAGAAGACCAACACACACTCAACCAATCAGAGACAGA 1527
Db      |||||
1244 TGGCCCGTGATGTGACGGGAACCTCAACACCTTCCGAACACCAAGAGCGGTGCAAGACGG 1185
Qy 1528 ATGCTTAAACATTTCATCTAAACACATACCTTGCTGGTGACTTTGGTTGGTTGACAAAAT 1587
Db      |||||
1184 ATGTTCAAATTTGAACCTCACCCGCGCTCGATCATGACTTTGGGAAGGTCACCAACGAG 1125
Qy 1588 GAA 1590
Db      |||
1124 GAA 1122

RESULT 12
US-09-246-320-2/c
; Sequence 2, Application US/09246320
; Patent No. 6251677

```

PECSIT.T 12

RESOLUTION 12  
IIS-09-246-320-2/c

US-09-248-320-2/C  
: Sequence 2. Application US/09246320; sequence 2, Application  
; Patent No. 6251677

100-443887-1

```

Db      1184  ATGTTCAAAATTGAACTACCCGCCGCTCTGGATCATGACTTTTGGGAAGGTCAACCAAGCAG 1122
Qy      1588  GAA 1590
        |||
Db      1124  GAA 1122

RESULT 13
US-09-546-738-2/c
; Sequence 2, Application US/09546738
; Patent No. 6387368
; GENERAL INFORMATION:
; APPLICANT: Trustees of University of Pennsylvania
; Wilson, James M.
; Kelley, William M.
; Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/546,738
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/246,320
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN.007PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-546-738-2

Query Match          5.4%; Score 109.4; DB 3; Length 4910;
Best Local Similarity 53.7%; Pred. No. 4.2e-21;
Matches 227; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

Qy      1168  GGC AAAAGAATACTGTTTATTTTCATGGACCCAGCCAGCAAAATCTATTATTGCA 1227
        |||
Db      1544  GGCAAGAGGAACACCATCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1485
        |||
Qy      1228  CAGCCCATAGCACAGCAGTGGCAATGTGGTTGCTATATGACGAGCAATGTAAACTTT 1287
        |||
Db      1484  GAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATAGAACTTT 1425
        |||
Qy      1288  CCATTTAATGACTCTACCAACAGAACTTGATTTGGGTAGAGAAGCTGGTAACTTTGGA 1347
        |||
Db      1424  CCCTTCAACGACTGTGTTCGACAGATGGTGTATCTGGTGGGAGGAGGAGATGACCGCC 1365
        |||
Qy      1348  CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTCGCAATTGATCAAAA 1407
        |||

```

Db	1364	AAGTCTGTGAGTGGCGGCAAAAGCCATTCTTCGGAGGAAGCAAGGTGCGCGTGGACCGAAA	1303
Qy	1408	GGAAAAGCGCAGCAAAACAGATTGGAACCAACACCAGTCTCATGTAGCACCACAAATGGAACATT	1467
Db	1304	TGCAAGTCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACTTCCAAACCAACATG	1245
Qy	1468	ACAGTGTGTGAGATAGGCTCGGAAGAAAGACACAGACACACTCAACCAATCAGAGACAGA	1527
Db	1244	TGCGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCAGCAGCGCGTTGCAAGACCCGG	1185
Qy	1528	ATGCTTAAACATTCATCTTAACACATACCTTCGCTGGTGACTTTTGGTTGGTTGCACAAAAT	1587
Db	1184	ATGTTCAATTTGAACTCACCGCGTCTGGATCATGACTTTGGAAAGGTCAACCAACGAG	1125
Qy	1588	GAA 1590	
Db	1124	GAA 1122	
RESULT 14			
US-09-438-268-1/c			
; Sequence 1, Application US/09438268			
; Patent No. 6491907			
; GENERAL INFORMATION:			
; APPLICANT: Rabinowitz, Joseph B.			
; APPLICANT: Samulski, Richard J			
; APPLICANT: Xiao, Weidong			
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING			
; TITLE OF INVENTION: THE SAME			
; FILE REFERENCE: 5470-186			
; CURRENT APPLICATION NUMBER: US/09/438,268			
; CURRENT FILING DATE: 1999-11-10			
; EARLIER APPLICATION NUMBER: 60/107,840			
; EARLIER FILING DATE: 1998-11-10			
; EARLIER APPLICATION NUMBER: 60/123,651			
; EARLIER FILING DATE: 1999-03-10			
; NUMBER OF SEQ ID NOS: 59			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 7214			
; TYPE: DNA			
; ORGANISM: Virus			
US-09-438-268-1			
Query Match 5.4%; Score 109.4; DB 4; Length 7214;			
Best Local Similarity 53.7%; Pred. No. 5.2e-21;			
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;			
Qy	1168	GGCAAAAGAATACTGTTTTATTTCATATGGACACGACGACGACAAATCTATTATTGCA	1227
Db	3899	GGCAAGAGGAACACCATCTGGCTGTTTGGCGCTGCAACTACCGGAAGACCAACATCGCG	3840
Qy	1228	CAAGCCATGACACAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCCATGTAAACTTT	1287
Db	3839	GAGGCCATGCCACACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT	3780
Qy	1288	CCATTTAATGACTGTACCAACAAGACTTCATTTTGGGTAGAAAGCTGGTAACTTTGGA	1347
Db	3779	CCCTTCAACACTGTGTGCAAGATGGTGATCTGGTGGAGGAGGAGATGACCGCC	3720
Qy	1348	CAGCAAGTAAACAGTTTAAAGCCATTGCTCTGGTCAAACTATTGCGATTGATCAAAA	1407
Db	3719	AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGGTGGACCCAGAA	3660
Qy	1408	GGAAAGCGACGAAAACAGATTGAACCAACACCACTCATCATGACCACAAATGAGAACATT	1467
Db	3659	TGCAAGTCTTCGGCCAGATAGACCCGACTCCCGTGATCGTCACTTCCAAACCAACATG	3600
Qy	1468	ACAGTGTGTGAGATAGGCTCGGAAGAAAGACACAGACACACTCAACCAATCAGAGACAGA	1527
Db	3599	TGGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCAAGCAGCGCGTTGCAAGACCGG	3540
Qy	1528	ATGCTTAAACATTCTTAACACATACCTTCGCTGGTGACTTTGGTTGGTTGACAAAAAT	1587

Db 3539 ATGTTCAAATTTGAACCTCACCGCGCTCTGGATCATGACTTTGGGAAGGTCACCAAGCAG 3480  
Qy 1588 GAA 1590  
Db 3479 GAA 3477

## RESULT 15

US-09-770-315-3  
; Sequence 3, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 7557  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-3

Query Match 5.4%; Score 109.4; DB 4; Length 7557;  
Best Local Similarity 53.7%; Pred. No. 5.3e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
Qy 1168 GGCAGAAAGAAATCTGTTTATTTTCATGACCGCCAGCAGCGAAATCTATTATGCA 1227  
Db 1321 GGCAGAGGAACACCATCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1380  
Qy 1228 CAAGCCATAGCACAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCATGTAACTTT 1287  
Db 1381 GAGGCCATAGCCACACTGTGCCCTTCTACGGTGCCTAACTGGACCAATGAGAACTTT 1440  
Qy 1288 CCATTTAATGACTGTACCAACAGAACTTGATTTGGGTAGAAGAAGCTGGTAACTTTTGA 1347  
Db 1441 CCCTTCAACGACTGTGCGACAAGATGGTGATCTGTGGGAGGAGGGAAGATGACCGCC 1500  
Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTGCTCTGGTCAAACTATTGCAATTGATCAAAA 1407  
Db 1501 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGAGGAAGCAAGGTGCGCGTGGACCAAAA 1560  
Qy 1408 GGAAGGCGAGCAAAACAGATTGAACCAACACCAAGTCATCATGACCAAAATGAGAACATT 1467  
Db 1561 TGCAGTCTCTGGCCCGAGATAGACCGACTCCCGTGATGTCATCTCCAAACCAACATG 1620  
Qy 1468 ACAGTGGTCAAGATAGGCTCGGAAGAAAGACAGAACACACTCAACCAATCAGAGACAGA 1527  
Db 1621 TGGCCCGTGATTGACGGGAACCTCAACGACCTTCAACACAGCAGCGGTTGCAAGACCGG 1680  
Qy 1528 ATGCTTAAACATTATCAACATACCTTGCCTGGTGACTTTGGTTGGTTGACAAAAAT 1587  
Db 1681 ATGTTCAAATTTGAACCTCACCCCGCTGATCATGACTTTGGGAAGGTCACCAAGCAG 1740  
Qy 1588 GAA 1590  
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:05:44  
Job time : 171.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds  
(without alignments)  
10821.780 Million cell updates/sec

Title: US-10-069-056-4  
Perfect score: 2019  
Sequence: 1 atggctggaagtcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	US-10-647-111A-1	Sequence 1, Appli
2	943	46.7	5075	US-10-361-002-3	Sequence 3, Appli
3	943	46.7	5075	US-10-361-004-3	Sequence 3, Appli
4	307.4	15.2	374	US-10-466-894-1276	Sequence 1276, Ap
5	285.8	14.2	421	US-10-466-894-430	Sequence 430, App
6	284.8	14.1	423	US-10-466-894-375	Sequence 375, App
7	284.8	14.1	464	US-10-466-894-370	Sequence 370, App
8	284.8	14.1	473	US-10-466-894-367	Sequence 367, App
9	284.8	14.1	466	US-10-466-894-358	Sequence 358, App
10	284.8	14.1	491	US-10-466-894-428	Sequence 428, App
11	284.2	14.1	420	US-10-466-894-438	Sequence 438, App
12	283.8	14.1	343	US-10-466-894-388	Sequence 388, App

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, Ap
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	115.4	5.7	1386	9	US-09-792-630-32	Sequence 32, Appl
40	115.4	5.7	1386	10	US-09-953-351-32	Sequence 32, Appl
41	115.4	5.7	1386	13	US-10-080-376-32	Sequence 38, Appl
42	115.4	5.7	1386	14	US-10-082-671-38	Sequence 32, Appl
43	115.4	5.7	1386	14	US-10-097-100-32	Sequence 32, Appl
44	115.4	5.7	1386	15	US-10-023-208-32	Sequence 28, Appl
45	113.8	5.6	1884	9	US-09-792-630-28	

ALIGNMENTS

RESULT 1

US-10-647-111A-1  
; Sequence 1, Application US/10647111A  
; Publication No. US20040209240A1  
; GENERAL INFORMATION:  
; APPLICANT: IGGO, RICHARD  
; APPLICANT: MALERBA, MADDALENA  
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES  
; FILE REFERENCE: 604-691  
; CURRENT APPLICATION NUMBER: US/10/647,111A  
; CURRENT FILING DATE: 2003-08-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 5121  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1  
; OTHER INFORMATION: with promoter P4 and left hairpin from WVM (pH1)  
US-10-647-111A-1

Query Match	87.7%	Score	1771;	DB	18;	Length	5121;
Best Local Similarity	92.3%	Pred.	No. 0;				
Matches	1864;	Conservative	0;	Mismatches	155;	Indels	0;
Gaps	0;						
QY	1	ATGGCTGGAAGTCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
Db	282	ATGGCTGGAAGTCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
QY	61	AGTAACCAAGAGTGTCTTCATTTTAAAAAAGAAATGTTCAACTGAATGAAAA	120				
Db	342	AGTAACCAAGAGTGTCTTCATTTTAAAAAAGAAATGTTCAACTGAATGAAAA	401				
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCGAGCTGAAATCTTTACAA	180				

Db	402	GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGAGGAGCAGAGCTGAAATCTTTACAA	461
Qy	181	CGAGGAGCGAAACTACTTCGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	240
Db	462	CGAGGAGCGAAACTACTTCGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	521
Qy	241	GAATGACCAAAAGCAAGTATTCAATTTTGATTCTTTGGTTAAAAATGTTTATTGAA	300
Db	522	GAATGACCAAAAGCAAGTATTCAATTTTGATTCTTTGGTTAAAAATGTTTATTGAA	581
Qy	301	GTGCTTAAACAAGAATATATTTCTGTGTGATGTTAAATTTGGTTTGTGCAACATGAATGG	360
Db	582	GTGCTTAAACAAGAATATATTTCTGTGTGATGTTAAATTTGGTTTGTGCAACATGAATGG	641
Qy	361	GGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTACTCAAGCT	420
Db	642	GGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTACTCAAGCT	701
Qy	421	CAAGGGAAATGCTGGAGAAGCAACTAAATGTTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480
Db	702	CAAGGGAAATGCTGGAGAAGCAACTAAATGTTTTACTTGGAGCAGATGGTTGGTAAACAGCC	761
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT	540
Db	762	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT	821
Qy	541	GAGTGGGTTACTCTACTTACTTATAGCATAAACCAACCAAAAGAGACTATACCAAGTGT	600
Db	822	GAGTGGGTTACTCTACTTACTTATAGCATAAACCAACCAAAAGAGACTATACCAAGTGT	881
Qy	601	GTTCTTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAAAGAAAAAAGACACTAGT	660
Db	882	GTTCTTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAAAGAAAAAAGACACTAGT	941
Qy	661	CCACCAAGAGACGAGGCTATTTTCTTTAGCAGTGACTCTGGCTTGGAAAACTAACTTTTAA	720
Db	942	CCACCAAGAGACGAGGCTATTTTCTTTAGCAGTGACTCTGGCTTGGAAAACTAACTTTTAA	1001
Qy	721	AAAGAGGCGAGCGCCATCTAGTGCAGCAACTATACATGATGACATGCGGCAGCAAAACG	780
Db	1002	AAAGAGGCGAGCGCCATCTAGTGCAGCAACTGTATACTGATGAGATGAACACAGAAACG	1061
Qy	781	GTTGAAACACAGTAAACCACTGGCGCAGGAACTAAAGCGCGCAGAAATTCAAAATTA	840
Db	1062	GTGAGACCAAGTGACCACTGCACAGGAGCTAAGCGCGCAGAAATTCAAAATAGAGAG	1121
Qy	841	GAAGTTGCTATTAACAATACTAAAGAGCTGGTGATTAAGAGTAACTCACAGAG	900
Db	1122	GAGGTCTCGATTAACAACCACTCAAGAGTTGGTACATATAAGAGTAACTCACAGAA	1181
Qy	901	GACTGGATGATGACGACAGACAGTTACATTAAGAAATGATGCTCAACAGAGTGGAGAA	960
Db	1182	GACTGGATGATGATGACGACAGACAGTTACATTAAGAAATGATGCTCAACAGAGTGGAGAA	1241
Qy	961	AACCTGCTGAAAAATAGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA	1020
Db	1242	AACCTGCTTAAAAATACACTAGAGATCTGTACACTGACTCTAGCAAGAACCAAAACAGCC	1301
Qy	1021	TTTGACTTAATTTTAAAAAGCTGAAACACAGCAAACTAAACCACTTTTCACTGCCTGAC	1080
Db	1302	TTTGACTTGAATCTGGAAAAAGCTGAAACACAGCAAACTAGCCAACTTTTCCATGGCTAGC	1361
Qy	1081	ACAGAACCTCGAGAAATTTTGTCTTTTATGCTGGTGAACCTATGTATGAAGTTTCCCATGCT	1140
Db	1362	ACCAAGACCTGTAGAAATCTTTGTGTGAGCATGGCTGGAACTATATTAAGTCTGCAATGCC	1421
Qy	1141	ATTTGCTGTGTTTTAAAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCATGGACCA	1200
Db	1422	ATCTGTTGTGTGTAATAGACAGGAGGCAAAAGGAAACACTGTGTCTTTTCAGGACCA	1481
Qy	1201	GCCAGCACAGGCAAACTATATTTATGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT	1260

Db	1482	GCACGACAGCGAAATCTATTATTGCAACAGCCATAGCACACGACGTTGGTAACTTTGGT	1541
Qy	1261	TGCTATAATGCAGCCAAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTGAAT	1320
Db	1542	TGTTACAATGCTGCCAATGTGAACCTTTTCCATTTAATGACTGTACCAACAAAACCTTTGAT	1601
Qy	1321	TGGGTAGAGAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTCCTCT	1380
Db	1502	TGGGTGGAAGAAGCTGGTAACTTTTGGCCAGCAAGTAAACCAATTTCAAGAGCTATTTGTTCT	1661
Qy	1381	GGTCAAACTATTCGCATTGTATCAAAAGGAAAAAGGACAGAAACAGATTGAACCAACACCA	1440
Db	1662	GGCCAAACCAATACGCATTGATCAAAAGGAAAGGACGAAACAGATTGACCAACACCA	1721
Qy	1441	GTCAATCATGACCACAAATAGAAACATTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA	1500
Db	1722	GTTTATTATGACCACCAACGAGAAACATTACCGTGGTTAGAATAGGCTGTGAGGAAAGACCA	1781
Qy	1501	GAACACACTCAACCAATCAGAGACAGNATGCTTTAAACATTTCATCTAACACATACCTTGCCT	1560
Db	1782	GAACACACTCAACCAATCAGAGACAGNATGCTCAACATTTCACCTGCACAGTACACTACCT	1841
Qy	1561	GGTCACCTTTGGTTTGGTTGTACAAAATAAGATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
Db	1842	GGTGACTTTGGTTTGGTTGGATAGACAGATGGCTCTGATCTGTGCTTGGTTGGTGAAG	1901
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGTATGCG	1680
Db	1902	AATGGTTACCAATCTACCATGGCTTGTGTTACTGTGCTAAATGGGGCAAAAGTTCTGTATGCG	1961
Qy	1681	TCAGAAACTGGGGGAGCCAAAGGTGCCAACTCCCTATAAATTTACTAGGTTCCGGCACGC	1740
Db	1962	TCAGAGACTTGGGGGAGCCGAAGCTAGACACTCCCTATAAATTCGGCTAGGTTCAATGCGC	2021
Qy	1741	TCACACTTCAGCACACCGAAAAGTACGCCCTCTCAGCCAGAACTATGCATCTCAACTT	1800
Db	2022	TCACACTCTGTACTCCGAGNATGACGCTCTCAGCCAAAACCTACGCTCTTACTCCACTT	2081
Qy	1801	GCATCGGACTCTCGAGGACCTGGCTTTTAGAGCCTTGGAGCACACAAATATCTCTGTTGCG	1860
Db	2082	GCATCGGACCTTGGGACCTTAGCTCTAGAGCCTTGGAGCACAGCAAAATACTCTCTGTTGCG	2141
Qy	1861	GGCACTCCGAAACCCAGAACACTGGGGAGCTGGTTTCCAAAGCCTGCCAAAGATGTTCAA	1920
Db	2142	GGCACTCGAGCAAGCCAAAACACTGGGGAGGCTGGTTTCCACAGCCTGCCAAGGTGTTCAA	2201
Qy	1921	CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCTTGGTGGCGAACCG	1980
Db	2202	CGGAGCCCAACTTGGTCCGAGATCGAGGCGGATTTGAGAGCTTGTCTTCACTCAAGAACAG	2261
Qy	1981	TTGAAGAAAGACTTCACGCGAGCCGCTGAACCTTGCACTAA	2019
Db	2262	TTGAGAGCGCACTTCAACGAGGAGCTGACCTCTGGACTAA	2300

```

RESULT 2
US-10-361-002-3
; Sequence 3, Application US/10361002
; Publication No. US20040170954A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gillemeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Arnistread David
; TITLE OF INVENTION: Pathogen Inactivation Assay
; FILE REFERENCE: CI-0043
; CURRENT APPLICATION NUMBER: US/10/361,002
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 5075

```

TYPE: DNA  
ORGANISM: Porcine parvovirus  
US-10-361-002-3

Query Match 46.7%; Score 943; DB 17; Length 5075;  
Best Local Similarity 70.5%; Pred. No. 9.3e-238;  
Matches 1306; Conservative 0; Mismatches 535; Indels 12; Gaps 3;

QY 4 GCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAAGT 63  
DB 298 GCGGGAACACCTTACTCGGAAGAGGTACTAAAAGCTTACCAACTGGCTTCAAGATAATGCT 357

QY 64 AACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAAAGAT 123  
DB 358 CAAAAAGAGCAATCTCTTATGTAATTTAAAAACAAAGATCAATCTAAATGGAAGAA 417

QY 124 ATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTTACAACGA 183  
DB 418 ATTGCTTGGATTAACCTACACAAGATACACAGATGCGGAATGATTAACCTACAAGA 477

QY 184 GGAGCGGAACCTACTTGGGACCAAGCGAGGACATGGAATGGGAACCAACAGTGGATGA 243  
DB 478 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAATCGACAG 537

QY 244 ATGACCAAAAAGCAAGTATTCATTTTGGTTTAAATGTTTAAATGTTTAAATGTTG 303  
DB 538 CTCACAAAACGGCAAGTACTGATTTTGGTCTCTTTTAAATAATGTTCTCTTTGAAGT 597

QY 304 CTTTACACAAAAGATATATTTCTGCTGATGTTTAAATGTTGTTGCAACATGAATGGGA 363  
DB 598 ATATTGCAAAAAGAACTTAAGTCAAGTACTGCTGTTTATACAGATGAATGATGTT 657

QY 364 AAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCTCA 423  
DB 658 CAAGTACTGGCTATCACTGCCATGTACTAGTGGGAAGGCTTACAAACAAGCAATG 717

QY 424 GGAATATGGTGAAGAGGCACTAATATTTTACTGGAGCAGATGGTTGGTAACAGCTGT 483  
DB 718 GGAATATGGTTCAGAAAAACAAATTAATTAATGAGTAGATGGTTAATATGCAATGC 777

QY 484 ATGTGCAACTACACAGCTGAAGAAATTAATAAGGAAATAGCAGAAACAATGAG 543  
DB 778 AAGTACTCTACACAGTTGAAGAAATTAATAAGGAAATAGCAGGATGGTGAG 837

QY 544 TGGGTTACTCTACTTACTTATAGCATAGCAACCAACCAACCAATATACAAATGACT 603  
DB 838 TGGGTATCGCTACTAACCCTACACTCACAAACAATTAATAACAATATACAAATGACT 897

QY 604 CTTTTTGGAAAAATGATGCTTACTATTTTAACTAATAAGAAATAGCACTAGTCCA 663  
DB 898 CATTTTGGAAATATGATGCTTACTACTTCTTAATAAATAAAGAAAGAACAACT 951

QY 664 CCAAGAGCGGAGCTTATTTCTTACAGTACTGCTGGCTGGAAACTTACTTTTAAAA 723  
DB 952 GAAAGAGAGCATGGATATATCTCAGCTCAGATCTCGGCTTATGCAAAATTTCTTAAA 1011

QY 724 GAAGGCGAGCGCCATCTAGTGAACAACTATACACTGATGACATGCGGCGAGAAACGGTT 783  
DB 1012 GAAGGCGAGAGACACTTAGTCACTACCTATTTACTGAGCAATTAACCTGAACTGTG 1071

QY 784 GAAACCAAGTAAACCACTGGCGAGGAACTAAGCGGCGCAGAAATCAAACTAAAAAGAA 843  
DB 1072 GAAACCAACGGTTACTACAGCTCAGGAAGCCAAAGAGGCGCAGATACAAACAAAAAGAA 1131

QY 844 GTTGCTTATTAATACTACACTTAAGAGCTGGTGCATTAAGAGTAACTCACCAGAGGAC 903  
DB 1132 GTAAGCATAAATGCACAATTAAGAGCTGGTTAATAAAGATGTACTAGCATAGAAGAC 1191

QY 904 TGGATGATGTCAGGCGAGACAGTATACATTTGAAATGATGCTCAACAGAGTGGAGAAAA 963  
DB 1192 TGGATGATGACATCCAGACAGTATATAGAAATGATGGCTCAACCGGAGAGAAAT 1251

QY 964 CTGCTGAAAAATACGTAGAGATTTGTACATAACTCTAGCCAGAACCAACAGCATTT 1023

DB 1252 TTAATCAAAAAATACACTAGAAATAACAACCTTTACTCTAGCAAGAACAAAAACAGCATAT 1311

QY 1024 GACTTAATTTTGAAGAAAGCTGAAACCAAGCAAACTAACCAAACTTTTCACTGCCTGCACACA 1083

DB 1312 GACTTAATACTTTGAAAGAGGCAAAACCAAGCATGCTACCAACATTTAATATTAGCAATACA 1371

QY 1084 AGAACCTGCAGAAATTTTGGCTTTTCATGGCTGGAACATATGTTAAAGTTTGGCAATGCTATT 1143

DB 1372 AGAACATGTAAATATTTAGCATGCAAAATTTGGAACATACATTAAGAGTCTGCCATGCTATA 1431

QY 1144 TCGTGTGTTTTTAAACAGACAGAGGAGCAAAAGAAATACTGTTTTTATTTTATGACACAGCC 1203

DB 1432 ACTTGTGTACTTAAACAGACAGAGGAGAAAGAAATACAAATTTCTATTTTATGGGCCAGCA 1491

QY 1204 AGCACAGGCAAAATCTATTATTGCAAGGCATAGCAACAAGCAGTTGGCAATGTTGGTTGC 1263

DB 1492 TCAACAGGAAAGATATAATTTGCTCAACATTTGCAAACTTAGTTGTAATGTTGGTTGC 1551

QY 1264 TATAATGCAAGCAATGTAAACCTTTCCATTTAATGACTGTACCAACAAGAACTTTGATTTGG 1323

DB 1552 TACAATGCAAGCAATGTGAACCTTTCCATTTAATGACTGTACAAATTAATAATG 1611

QY 1324 GTAGAGAAAGCTGGTAACTTTGGACAGCAAGTAAACCAAGTTTAAAGCATTTCCTCTGCT 1383

DB 1612 ATTGAAGAGAGCAGGAACTTTCTTAACCAAGTAAACCAATTTCAAGCCATATGTTCAAGT 1671

QY 1384 CAAACTATTTCGCAATTCATCAAAAGGAAAGGCAAGCAAGACAGATTGAAACCAACCAAGTC 1443

DB 1672 CAAACAATTAGAATTGACCAAAAGGTTAAGGAAGCAACAATTTGAACCACTCTCTGTA 1731

QY 1444 ATCATGACCAAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTGCGAAAGAAAGACAGAA 1503

DB 1732 ATAATGACTACAAATGAGACATAACTTAAAGTTTAAAGTAGGATGCGAGGAAAGACCAAG 1791

QY 1504 CACACTCAACCAATCAGACAGCAAGTCTTAACTTCACTCACTACACATACCTTTGCTCTGCT 1563

DB 1792 CATACAACCAATTAAGACAGCAAGTGTAAACCAATTAACCAACCAAGAACTGCCAGT 1851

QY 1564 GACTTTGGTTTGGTTGACCAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAGAAAT 1623

DB 1852 GATTTTGGACTTTTAGAAGAACTGAATGGCCCATTAATATGCTTTGGTTGGTAAAGAA 1911

QY 1624 GGTTCACATCTACCATGGCAAGCTACTGTCTTAATTTGGGGCAAGTTCTCTGATTTGCTCA 1683

DB 1912 GGTTCACAGCAACATTTGCTAGCTATATGCTATGCTATTTGGGAAATGTACCTGATTTGCTCA 1971

QY 1684 GAAACTGGGCGAGCCAAAGGTGCAACTCCCTATAAATTTACTAGTTTGGGCAAGCTCA 1743

DB 1972 GAAATATGGAGGAGCCAAATTTGCAACCCCAATTAATACACCA---ACAGACTCTCAG 2028

QY 1744 CCATTCACACACCGGAAAAAGTAGCCTCTCAGCCAGAACTATATGCACTTAACCTCCACTTGCA 1803

DB 2029 ATTTCACATCAGTGAAACTTTGCCAGCGGACCAACACTACGACGCACTCCCAATACAG 2088

QY 1804 TCGGATCTCGA---GGACCTGGCTTTAGAGCTTTGAGCAGCACCAAAATATCTC 1853

DB 2089 GAGGACCTGGATTTAGCTTTAGCTTTGGAGCCGTGGAGCGGAGCCAAACACACC 2141

RESULT 3  
US-10-361-004-3  
; Sequence 3, Application US/10361004  
; Publication No. US20040170981A1  
; GENERAL INFORMATION:  
; APPLICANT: Clearant, Inc.  
; APPLICANT: McKenney, Keith  
; APPLICANT: Gillmeister, Lidja  
; APPLICANT: Marlowe, Kristina  
; APPLICANT: Armistead, David  
; TITLE OF INVENTION: Real-Time Polymerase Chain Reaction Using Large Target Amplicons  
; FILE REFERENCE: CI-0042  
; CURRENT APPLICATION NUMBER: US/10/361,004

; CURRENT FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 5075  
; TYPE: DNA  
; ORGANISM: Porcine parvovirus  
US-10-361-004-3

Query Match 46.7%; Score 943; DB 17; Length 5075;  
Best Local Similarity 70.5%; Pred. No. 9.3e-238;  
Matches 1306; Conservative 0; Mismatches 535; Indels 12; Gaps 3;  
4 GCTGGAAATCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAAAGT 63  
298 GCGGGAACACTTACTCGGAAGAGGTACTTAAAGCTACCACTGGCTTCAAGATATGCT 357  
64 AACGAGAAGTGTCTCATTTGTTTAAATGAAATGTTCAACTGAATGAAAGAT 123  
358 CAAAAAGAAGCAATCTCTTATGTTTAAACACAAAAAGTCAATCTAAATGAAAAAGAA 417  
124 ATCGGATGGAATGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACACGA 183  
418 ATTGCTTGGAAATCACTACAAAGAGATACACAGATGCGGAATGATAACCTTACAAGA 477  
184 GGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGATGAA 243  
478 GGAGCAGAAACATCATCGGACAGGCAACACAGACATGGAATGGGAATCAGNAATCGACAGC 537  
244 ATGACCAAAAGCAAGTATTCATTTTGTGATCTTTTGGTTTAAATAATGTTTATTTGAAGTG 303  
538 CTCACAAAACGGCAAGTACTGATTTTGTGCTCTCTTTGTTAAAAAATGCTCTTTTGAAGGT 597  
304 CTTTAAACACAAAGATATATTTCTGTTGATGTTTAAATTTGTTGTCGAACATGAATCGGGA 363  
598 ATATTGCAAAAGAACCTTAAGTCCAAGTACTGCTGTTTATCAGCATGAACATGTT 657  
364 AAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTCAAGCTCAA 423  
658 CAAGATACTGGCTATCACTGCCATGTACTTAGTGGTGGAAAGGCTTACAAACAGCAATG 717  
424 GGAATATGGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCTGT 483  
718 GGAATATGGTTTCAAGAAACAAATTTAAACAAATTTATGAGTAGATGTTTAAATGCAATGC 777  
484 AATGTGCAACTAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAAGCAATGAG 543  
778 AAGTACTCTTAACACAGTTGAAAGAAATTAATTAAGGAATTAGCAGGATGGTGAG 837  
544 TGGGTTACTCTTACTTATAGCATAAGCAACCAACCAACCAACCAATATACAAAAATGACT 603  
838 TGGGTTATCGCTACTAACCCTACACTCACAAACCAACCAACCAATATACAAAAATGACT 897  
604 CTTTTTGGAAACATGATGTTCTTACTTATTTTAACTTAAAGAAATAGACTAGTCCA 663  
898 CATTTTGGAAATATGATGTTCTTACTTCTTAAATAAAAAAGAAAGCAACT----- 951  
664 CCAAGAGACGGAGGCTATTTCTTACAGTGTGCTGGCTGGGAAACTTAACTTTTAAAA 723  
952 GAAAGAGAGCATGGATATTAATCTCAGCTCAGATCTCGGCTTCAATCAAAATTTCTTAAAA 1011  
724 GAAGGCGAGCGGCATCTAGTGAGCAAACTATACATGATGACATGCGGCGCAGAAACGGTT 783  
1012 GAAGGCGAGAGACACTTAGTCACTACCTATTTTACTTGAAGCAATAAACCTGAAACTGTG 1071  
784 GAAACACAGTAACTAGTGGCAGGAAACCTAAGCGGCGGAGAAATTCAACTTAAAGAA 843  
1072 GAAACACAGTAACTAGTGGCAGGAAACCTAAGCGGCGGAGAAATTCAACTTAAAGAA 1131  
844 GTTGTCTTAAATACTACCTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAGGAC 903  
1132 GTNAGCATTAATGCAATAGAGACTTGGTTTAAATTAAGATGTTACTAGCATAGAGAC 1191

## RESULT 4

US-10-466-894-1276  
; Sequence 1276, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telesman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius

904 TGGATGATGATGCGAGCCAGACAGTGTACATTTGAAATGATGGCTCAACAGGTTGGAGAAAAC 963  
1192 TGGATGATGACAGATCCAGACAGTTATATAGAAATGATGGCTCAACAGGAGGAGAAAAT 1251  
964 CTGCTGAAAAATACGCTAGAGATTTGTACACTTCTAGCCAGAAACCAAAACAGCATTT 1023  
1252 TTAATCAAAAATACACTAGAAAATAACAACCTTTACTCTAGCAAGAACAAAAACAGCATAT 1311  
1024 GACTTAATTTTAGAAAAAGCTGAAACCCAGCAACCACTAACTTTTCACTGCTGCACACA 1083  
1312 GACTTAATTTTAGAAAAAGCAAAACCAAGCATGCTTACCAACATTTTAATTTAGCAATACA 1371  
1084 AGAACCTGCAGAAATTTTCTTTTCATGCTGGAACATATGTTAAAGTTTGGCCATGCTATT 1143  
1372 AGAACATGTAATAATTTTCAAGCATGCAATTTGGAACCTACATTTAAAGTCTGCCATGCTATA 1431  
1144 TGTGTTGTTTTTAAACAGACAGAGGCAAAAGAAATACATGTTTATTTATTCATGACAGCC 1203  
1432 ACTTGTGTACTAAACAGACAGAGGAGAAAGAAATACAAATTTCTATTTTCATGGCCAGCA 1491  
1204 AGCACAGGCAAAATCTATTATTGCAAGCCATAGCACACAGCTTGGCAATGTTGGTTGC 1263  
1492 TCAACAGGAAAGATATAATTTGCTCAACATTTGCAAACTTAGTTGGTAATGTTGGTTGC 1551  
1264 TATAATGACAGCAATGTTAAACTTTCCATTTAATGACTGTACCAACAGAACTTTGATTTGG 1323  
1552 TACAATGACAGCAATGTTGAACCTTTCCATTTAATGACTGTACAAATAAATAAATTAATATGG 1611  
1324 GTAGAAGAAGCTGTTAACTTTTGGACAGCAAGTAAACAGCTTTTAAAGCCATTTGCTCTGTT 1383  
1612 ATTGAAGAAGCAGGAAACTTCTTAAACCAAGTAAACCAATTTCAAAGCCATATGTTTCAGGT 1671  
1384 CAAACTATTTCGCAATGATCAAAAAGGAAAGCAGCAACAGATTTGAACCAACAGCTC 1443  
1672 CAACCAATTTAGAAATTTGACCAAAAAGGTAAGAGCAAACTTTGAACCACTCTCTGTA 1731  
1444 ATCATGACCAAAATGAGAACATTAACAGTGTGAGATAGGCTGCGAAGAAAGACAGAA 1503  
1732 ATAAATGACTTACAAATGAGACATAACTAAAGTTAGATAGGATGCGAGGAAAGACAGAA 1791  
1504 CACACTCAACCAATCAGACAGCAAGTGTAACTTCACTTAAACATACATACCTTGGCTGTT 1563  
1792 CATAACACCAATTAAGACAGCAAGTGTAAACATAAAACCTTAAACAGAAACTGCGCAGGT 1851  
1564 GACTTTGGTTTGGTTGACAAAATGAAATGGGCCCATGATTTGCTTGGTTGGTAAAGAAAT 1623  
1852 GATTTTGGACTTTTAGAAGAACTGATGGCCCACTAATATGTTGTTGGTAAAGAA 1911  
1624 GGTACCAATCTACCATGCAAGCTTACTGTGCTTAAATGGGGCAAGTTCCTGATTTGGTCA 1683  
1912 GGTACCAAGCAACAAATGGCTAGCTATATGCTATGCTTGGGAAATGTTACCTGATTTGGTCA 1971  
1684 GAAACTGGGCGGAGCCAAAGTGGCCAACTCTTAAATAATTTACTAGTTTGGGCACTCA 1743  
1972 GAAATGGGAGGAGCCAAAATGCAAAACCCCAATTAATATACACCA---ACAGACTCTCAG 2028  
1744 CCATTCACGACACCCGAAAGTAGCGCTCTCAGCCAGAACTATGCTACTAATCTCCACTTGCA 1803  
2029 ATTTCCACATCAGTGAAGAACTTGGCAGCGGACAACTACGACGCAACTCCATACAG 2088  
1804 TCGGATCTCGA---GGACCTGGCTTTTAGAGCTTTGGAGCACACCAAAATCTCC 1853  
2089 GAGGACCTGGATTTAGCTTTAGCTTTGGAGCGGTGGAGCGGAGCCAAACACACC 2141



```
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-1276

Query Match      15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 1.7e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 174 TTTTACACGAGGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTACACGAGGCGGAGACCACTTTGGGACCAAGCGAGGACATGGAATGGGAGCGC 60

QY 234 AGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGTATCTTTGTTAAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTCTTTTGTATCTTTGTTAAAGAGTGT 120

QY 294 ATTGGAAGTCTTAACACAAAGATATATTTCTCTGTGTATGTTAAATGGTTGCAACA 353
Db 121 GTTTGAAGTCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTTCTGTTGTCAGCA 180

QY 354 TGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGCACTTTAG 413
Db 181 TGAATGGGAAAGACCAAGGCTGGCACTGTCTGCTGATTTGGAGGACAGCACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGGTGGAGAAAGCAACTAAATGTTTACTGGAGCAGATGGTTGGT 473
Db 241 TCAACCTCAAGGAAATGGTGGAGAAAGCAAGTAAATGTTGCTAGTANATGGTTGGN 300

QY 474 AACAGCTGTAAATGTCGAACCTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGA 533
Db 301 GACTGCTGNAATGTNCAACTAACACCAAGCTGAAAGAAATTAACCTGNNAGAAATAGCAGA 360

QY 534 AGACAATGAGTGG 546
Db 361 GGACAGTGANNGG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US2004024167A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
```

```
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 276, 405, 417, 419, 421
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-430

Query Match      14.2%; Score 285.8; DB 18; Length 421;
Best Local Similarity 87.6%; Pred. No. 9.4e-65;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAACCA 232
Db 59 CTTTACACGAGGAGCGGAGACCACTTTGGGACCAAGCGAGGACATGGAATGGGAGCG 118

QY 233 CAGTGTGTAATGACCAAAAGCAAGTATTCTTTTGTATCTTTGTTAAAAAATGTT 292
Db 119 CAGTGTGATGACATGACCAAAAGCAAGTATTCTTTTGTATCTTTGTTTAAAGAGT 178

QY 293 TATTTGAAGTGTCTTAACACAAAGATATATTTCTCTGTGTATGTTAAATGGTTTGTGCAAC 352
Db 179 TGTTTGAAGTGTCTCAGCACAAAGACATAGCTCTTAGTAAATGTTACTTGTTCGTGCGC 238

QY 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGACTTTTA 412
Db 239 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCTGCTGATTTGGAGGCAAGCACTTTA 298

QY 413 GTCAAGCTCAAGGAAATGGTGGAGAAAGCAACTAAATGTTTACTTGAGCAGATGGTTGG 472
Db 299 GTCAAGCTCAAGGAAATGGTGGAGAAAGCAAGTAAATGTTGTAAGTAGATGGTTGG 358

QY 473 TAAACGCTGTATGTGTCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAAT 527
Db 359 TGAAGTGTATGTTCAACTTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAAT 413

RESULT 6
US-10-466-894-375
; Sequence 375, Application US/10466894
; Publication No. US2004024167A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```

Query Match 14.1%; Score 284.8; DB 18; Length 423;  
Best Local Similarity 89.2%; Pred. No. 1.7e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACAACGAGGCGGAAACTACTTTGGGACCAAGCGGAGCATGGAAATGGGAACCA 232  
DB 71 CTTTACAACGAGGCGGAGACCACTTTGGGACCAAGCGGAGCATGGAAATGGGAGCG 130

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTCTGGTGAATGTTTAAAGATGTT 292  
DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTCTTCTGGTGAATGTTTAAAGATGTT 190

QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTCTGGTGAATGTTTAAAGATGTT 352  
DB 191 TGTTCGCTGCTCAGCACAAAGATATATTTCTCTGGTGAATGTTTAAAGATGTT 250

QY 353 ATGAATGGGAAAGCAAGGCTGGCACTGCCATGCTAATGTTTAAAGATGTTTAA 412  
DB 251 ATGAATGGGAAAGCAAGGCTGGCACTGCCATGCTAATGTTTAAAGATGTTTAA 310

QY 413 GTCAAGCTCAAGGGAATGTTGAGAGGCAACTAAATGTTTAAAGATGTTTAA 472  
DB 311 GTCAAGCTCAAGGGAATGTTGAGAGGCAACTAAATGTTTAAAGATGTTTAA 370

QY 473 TAACAGCCTGTAATGTGCAACTAACACCAAGCTGAAAGATTTAA 516  
DB 371 TGACTGCCTGTAATGTTCACCTAACACCAAGCTGAAAGATTTAA 414

RESULT 7  
US-10-466-894-370  
; Sequence 370, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Teleman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466, 894  
; PRIOR FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 370  
; LENGTH: 464  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 414\_435, 459, 462, 464  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-370

Query Match 14.1%; Score 284.8; DB 18; Length 464;  
Best Local Similarity 89.2%; Pred. No. 1.7e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACAACGAGGCGGAAACTACTTTGGGACCAAGCGGAGCATGGAAATGGGAACCA 232  
DB 55 CTTTACAACGAGGCGGAGACCACTTTGGGACCAAGCGGAGCATGGAAATGGGAGCG 114

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTCTGGTGAATGTTTAAAGATGTT 292  
DB 115 CAGTGGATGACATGACCAAAAGCAAGTATTCTTCTGGTGAATGTTTAAAGATGTT 174

QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTCTGGTGAATGTTTAAAGATGTTTAA 352  
DB 175 TGTTCGCTGCTCAGCACAAAGATATATTTCTCTGGTGAATGTTTAAAGATGTTTAA 234

QY 353 ATGAATGGGAAAGCAAGGCTGGCACTGCCATGCTAATGTTTAAAGATGTTTAA 412  
DB 235 ATGAATGGGAAAGCAAGGCTGGCACTGCCATGCTAATGTTTAAAGATGTTTAA 294

QY 413 GTCAAGCTCAAGGGAATGTTGAGAGGCAACTAAATGTTTAAAGATGTTTAA 472  
DB 295 GTCAAGCTCAAGGGAATGTTGAGAGGCAACTAAATGTTTAAAGATGTTTAA 354

QY 473 TAACAGCCTGTAATGTGCAACTAACACCAAGCTGAAAGATTTAA 516  
DB 355 TGACTGCCTGTAATGTTCACCTAACACCAAGCTGAAAGATTTAA 398

RESULT 8  
US-10-466-894-367  
; Sequence 367, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Teleman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466, 894  
; PRIOR FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 367  
; LENGTH: 473  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 10, 431, 447, 467, 468, 473  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-367

Query Match 14.1%; Score 284.8; DB 18; Length 473;  
Best Local Similarity 89.2%; Pred. No. 1.9e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACAACGAGGCGGAAACTACTTTGGGACCAAGCGGAGCATGGAAATGGGAACCA 232  
DB 72 CTTTACAACGAGGCGGAGACCACTTTGGGACCAAGCGGAGCATGGAAATGGGAGCG 131

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTCTGGTGAATGTTTAAAGATGTT 292  
DB 132 CAGTGGATGACATGACCAAAAGCAAGTATTCTTCTGGTGAATGTTTAAAGATGTT 191

QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTCTGGTGAATGTTTAAAGATGTTTAA 352  
DB 192 TGTTCGCTGCTCAGCACAAAGATATATTTCTCTGGTGAATGTTTAAAGATGTTTAA 251

QY 353 ATGAATGGGAAAGCAAGGCTGGCACTGCCATGCTAATGTTTAAAGATGTTTAA 412  
DB 252 ATGAATGGGAAAGCAAGGCTGGCACTGCCATGCTAATGTTTAAAGATGTTTAA 311

QY 413 GTCAAGCTCAAGGGAATGTTGAGAGGCAACTAAATGTTTAAAGATGTTTAA 472  
DB 312 GTCAAGCTCAAGGGAATGTTGAGAGGCAACTAAATGTTTAAAGATGTTTAA 371

QY 473 TAACAGCCTGTAATGTGCAACTAACACCAAGCTGAAAGATTTAA 516

```
Db 372 TGACTGCTGTATGTTCAACTAACACACAGCTGAAAAA 415
|||||
RESULT 9
US-10-466-894-358
; Sequence 358, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 413, 428, 435, 448, 457, 462
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;
Best Local Similarity 89.2%; Pred. No. 1.9e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232
Db 54 CTTTACACGAGGAGCGGAGACCACTTTGGACCAAGCGAGGACATGGAATGGGAGCG 113
QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAAATGTT 292
Db 114 CAGTGAATGACATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAGAGTGT 173
QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTGTGTATGTTAAATGTTTGTGCAAC 352
Db 174 TGTTGAAGTCTCAGCACAAAGACATAGCTCTCTAGTAATGTTACTTGTGCGAGC 233
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412
Db 234 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 293
QY 413 GTCAAGCTCAAGGAAAATGTTGAGAGGCAACTAAATGTTTACTGTGAGCAGATGTTGG 472
Db 294 GTCAAGCTCAAGGAAAATGTTGAGAGGCAACTAAATGTTTACTGTGAGTATGTTGGTGG 353
QY 473 TAACAGCTGTATGTGCACTTAACACCACTTAACACCACTTAACCACTTAACCACTTAAC 516
Db 354 TGACTGCTGTATGTTCACTTAACACCACTTAACACCACTTAACCACTTAACCACTTAAC 397

RESULT 10
US-10-466-894-428
; Sequence 428, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 12, 421, 441
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;
Best Local Similarity 89.2%; Pred. No. 1.9e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232
Db 62 CTTTACACGAGGAGCGGAGACCACTTTGGACCAAGCGAGGACATGGAATGGGAGCG 121
QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAAATGTT 292
Db 122 CAGTGAATGACATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAGAGTGT 181
QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTGTGTATGTTAAATGTTTGTGCAAC 352
Db 182 TGTTGAAGTCTCAGCACAAAGACATAGCTCTCTAGTAATGTTACTTGTGCGAGC 241
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412
Db 242 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 301
QY 413 GTCAAGCTCAAGGAAAATGTTGAGAGGCAACTAAATGTTTACTGTGAGCAGATGTTGG 472
Db 302 GTCAAGCTCAAGGAAAATGTTGAGAGGCAACTAAATGTTTACTGTGAGTATGTTGGTGG 361
QY 473 TAACAGCTGTATGTGCACTTAACACCACTTAACACCACTTAACCACTTAACCACTTAAC 516
Db 362 TGACTGCTGTATGTTCACTTAACACCACTTAACACCACTTAACCACTTAACCACTTAAC 405

RESULT 11
US-10-466-894-438
; Sequence 438, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 438
```

```

; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 418
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-438

Query Match
Best Local Similarity 14.1%; Score 284.2; DB 18; Length 420;
Matches 310; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 173 CTTTCAACGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAACCA 232
Db 64 CTTTCAACGAGGAGCGGAGGACCTTGGGACCAAGCGGAGGACATGGAATGGGAGCG 123

QY 233 CAGTGAATGAATGACAAAGCAAGTATTCATTTTGAATCTTTGGTTTAAAAATGTT 292
Db 124 CAGTGAATGACATGACCAAGCAAGTATTTATTTTGAATCTTTGGCTTAAGAGTGT 183

QY 293 TATTTCAAGTCTTAACACAAAGCAATATATTTCTCGTGATGTTAATTTGTTGCAAC 352
Db 184 TGTTTGAAGTCTTCAGACAAAGCAATAGCTCTCTAGTAATGTTACTTGTGTTGTCAGC 243

QY 353 ATGAATGGGAAAGCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGGAAAGGACTTTA 412
Db 244 ATGAATGGGAAAGCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGGAAAGGACTTTA 303

QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCACTAAATGTTTACTGGAGGAGATGTTGG 472
Db 304 GTCAAGCTCAAGGAAATGGTGAGAGGCACTAAATGTTTACTGGAGGATGTTGGTGG 363

QY 473 TACAGCCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAA 525
Db 364 TGACTGCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAA 416

RESULT 12
US-10-466-894-388
; Sequence 388, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Subini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 388
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-388

Query Match
Best Local Similarity 14.1%; Score 283.8; DB 18; Length 343;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGAGCGGAGGACCTTGGGACCAAGCGGAGGACATGGAATGGGAGCGC 60

QY 234 AGTGAATGAATGACCAAGCAAGTATTCATTTTGAATCTTTGTTTAAAAATGTTT 293
Db 61 AGTGAATGACATGACCAAGCAAGTATTTATTTTGAATCTTTGTTTAAAAAGTGT 120

QY 294 ATTTGAAGTCTTAACACAAAGCAATATTTCTCGTGATGTTAATTTGTTTGTCAACA 353
Db 121 GTTTGAAGTCTTCAGACAAAGCAATAGCTCTCTAGTAATGTTACTTGTGTTGTCAGCA 180

QY 354 TGAATGGGAAAGCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGGCAAGGACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGGTGAGAGGCACTAAATGTTTACTGGAGGAGATGTTGGT 473
Db 241 TCAAGCTCAAGGAAATGGTGAGAGGCACTAAATGTTTACTGGAGTAGATGTTGGT 300

QY 474 AACAGCCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTA 516
Db 301 GACTGCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTA 343

RESULT 13
US-10-466-894-390
; Sequence 390, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Subini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-390

Query Match
Best Local Similarity 14.1%; Score 283.8; DB 18; Length 343;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGAGCGGAGGACCTTGGGACCAAGCGGAGGACATGGAATGGGAGCGC 60

QY 234 AGTGAATGAATGACCAAGCAAGTATTCATTTTGAATCTTTGTTTAAAAATGTTT 293
Db 61 AGTGAATGACATGACCAAGCAAGTATTTATTTTGAATCTTTGTTTAAAAAGTGT 120

QY 294 ATTTGAAGTCTTAACACAAAGCAATATTTCTCGTGATGTTAATTTGTTTGTCAACA 353
Db 121 GTTTGAAGTCTTCAGACAAAGCAATAGCTCTCTAGTAATGTTACTTGTGTTGTCAGCA 180

QY 354 TGAATGGGAAAGCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGGCAAGGACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGGTGAGAGGCACTAAATGTTTACTGGAGGAGATGTTGGT 473
Db 241 TCAAGCTCAAGGAAATGGTGAGAGGCACTAAATGTTTACTGGAGTAGATGTTGGT 300

QY 474 AACAGCCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTA 516
Db 301 GACTGCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTA 343
```

Db 301 GACTGCTGTATGTTCAACTTAACACGCTGAAAAA 343

## RESULT 14

US-10-466-894-433  
; Sequence 433, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 433  
; LENGTH: 403  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 218  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-433

Query Match 14.1%; Score 283.8; DB 18; Length 403;  
Best Local Similarity 89.0%; Pred. No. 3.1e-64;  
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 173 CTTTACAACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCA 232  
Db 56 CTTTACAACGAGGAGCGGAGCCACTTTGGGACCAAGCGAGGACATGGAAATGGGAGCG 115  
QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCATTTTGGTAAATGTT 292  
Db 116 CAGTGAATGACATGACCAAAAGCAAGTATTTTGGTAAATGTT 175  
QY 293 TATTTGAAGTCTTAACACAAAGAAATATTTCTGCTGATGTTAAATGTTGCAAC 352  
Db 176 TGTTCGAAGTCTGACGACAAAGAAATAGCTCCTAGTAAATGTTGCTGAGC 235  
QY 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTTAATGGAGGAAAGGACTTTA 412  
Db 236 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCTGCTGATGGAGGCAAGGACTTTA 295  
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
Db 296 GTCAAGCTCAAGGAAATGGTGAGAGGCAAGGCACTAAATGTTTACTGGAGTAGATGTTGG 355  
QY 473 TACAGCCTGTATGTGCAACTTAACACAGCTGAAGAAATTA 516  
Db 356 TGACTGCTGTATGTTCACACTTAACACAGCTGAAGAAATTA 399

## RESULT 15

US-10-466-894-369  
; Sequence 369, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 369  
; LENGTH: 497  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 448-475  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-369

Query Match 14.0%; Score 283.2; DB 18; Length 497;  
Best Local Similarity 89.0%; Pred. No. 5e-64;  
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 173 CTTTACAACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCA 232  
Db 68 CTTTACAACGAGGAGCGGAGCCACTTTGGGACCAAGCGAGGACATGGAAATGGGAGCG 127  
QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCATTTTGGTAAATGTT 292  
Db 128 CAGTGAATGACATGACCAAAAGCAAGTATTTTGGTAAATGTT 187  
QY 293 TATTTGAAGTCTTAACACAAAGAAATATTTCTGCTGATGTTAAATGTTGCAAC 352  
Db 188 TGTTCGAAGTCTGACGACCAAGAAATAGCTCCTAGTAAATGTTGCTGAGC 247  
QY 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTTAATGGAGGAAAGGACTTTA 412  
Db 248 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCTGCTGATGGAGGCAAGGACTTTA 307  
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
Db 308 GTCAAGCTCAAGGAAATGGTGAGAGGCAAGGCACTAAATGTTTACTGGAGTAGATGTTGG 367  
QY 473 TACAGCCTGTATGTGCAACTTAACACAGCTGAAGAAATTA 516  
Db 368 TGACTGCTGTATGTTCACACTTAACACAGCTGAAGAAATTA 411

Search completed: January 23, 2005, 14:35:09  
Job time : 1076 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:18:38 ; Search time 99 Seconds  
(without alignments)  
2435.011 Million cell updates/sec

Title: US-10-069-056-6  
Perfect score: 3638  
Sequence: 1 MAGNAYDEVILGNTWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3628	99.7	672	4	AAU72704 Parvoviru
2	3625	99.6	672	4	AAU72702 Parvoviru
3	3620	99.5	672	4	AAU72708 Parvoviru
4	3620	99.5	672	4	AAU72710 Parvoviru
5	3620	99.5	672	4	AAU72706 Parvoviru
6	2373	65.2	690	1	AAU40306 Sequence
7	455.5	12.5	550	3	AAU58164 Adeno ass
8	455.5	12.5	550	3	AAU11409 Adeno-ass
9	455.5	12.5	610	3	AAU58159 Adeno ass
10	455.5	12.5	610	4	AAU97720 Rep prote
11	455.5	12.5	610	5	AAU11404 Adeno-ass
12	455.5	12.5	610	5	AAU22887 Adeno-ass
13	455.5	12.5	610	5	AAU28644 Adeno-ass
14	455.5	12.5	610	5	AAU26940 Adeno ass
15	455.5	12.5	610	6	AAU64865 Rep prote
16	455.5	12.5	610	6	AAU64760 Adeno ass
17	455.5	12.5	610	6	AAU43398 Adeno-ass
18	455.5	12.5	610	7	AAU40280 Adeno-ass
19	455.5	12.5	610	8	AAU58895 Predeterm
20	452.5	12.4	626	2	AAU85384 Barbary d
21	452.5	12.4	626	4	AAU97724 NS1 prote
22	452.5	12.4	626	5	AAU22891 Muscovy d
23	452.5	12.4	626	5	AAU28648 Muscovy d
24	452.5	12.4	626	5	AAU26944 Muscovy d
25	452.5	12.4	626	6	AAU64869 NS1 prote

26	452.5	12.4	626	6	AAU64764 Parvoviru
27	452.5	12.4	626	6	AAU43402 Muscovy d
28	452.5	12.4	626	7	AAU40288 Muscovy d
29	452	12.4	627	4	AAU97722 Rep prote
30	452	12.4	627	5	AAU22889 Barbarye
31	452	12.4	627	5	AAU28646 Barbarye
32	452	12.4	627	5	AAU26942 Barbarye
33	452	12.4	627	6	AAU64867 Rep prote
34	452	12.4	627	6	AAU64762 Parvoviru
35	452	12.4	627	6	AAU43400 Barbarye
36	452	12.4	627	7	AAU40284 Barbarye
37	445.5	12.2	625	7	AAU62760 Adeno ass
38	437.5	12.0	536	6	AAU46311 AAV4 Rep
39	437.5	12.0	536	6	AAU73938 Adeno-ass
40	437.5	12.0	623	2	AAU46307 AAV4 Rep
41	437.5	12.0	623	4	AAU46312 AAV4 Rep
42	437.5	12.0	623	4	AAU97712 Rep prote
43	437.5	12.0	623	5	AAU28636 Adeno-ass
44	437.5	12.0	623	5	AAU26932 Adeno ass
45	437.5	12.0	623	6	AAU64857 Rep prote

## ALIGNMENTS

RESULT 1  
AAU72704  
ID AAU72704 standard; protein; 672 AA.

XX AAU72704;  
XX AC  
XX AAU72704;  
DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX  
XX Parvovirus non-structure protein 1 (NS1) variant (S283A).  
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; variant.  
XX  
XX Parvovirus.  
OS Synthetic.  
OS  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 283  
FT /note= "Wild type Ser substituted with Ala"  
XX  
XX EP1077260-Al.  
XX  
XX 21-FEB-2001.  
XX  
XX 13-AUG-1999; 99EP-00115161.  
XX  
XX 13-AUG-1999; 99EP-00115161.  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
XX Nueesch J, Rommelaere J;  
XX WPI: 2001-212717/22.  
XX N-PSDB; AAD02799.  
XX  
XX Novel parvovirus non-structure protein variant, useful for treating  
XX tumoral diseases, has a shifted equilibrium between DNA replication and  
XX transcription activities, and cytotoxic activity.  
XX  
XX Claim 6; Page 14-16; 41pp; English.  
XX  
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant  
XX (S283A). The invention relates to the variants of the parvovirus non-  
XX structure protein (NS1) having a shifted equilibrium between the DNA  
XX replication and transcription activities, and the cytotoxicity activity.  
XX These variants are useful as vectors for treating tumoural diseases. The  
XX variant DNAs are useful for gene therapy. (Updated on 06-AUG-





```

Db 541 NGQSTWASYCAKWKVPDSENWAEKVPPTPINLLGSARSPETTPKSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETONTGEGSKACODGOLSPTWSEIEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAETONTGEGSKACODGOLSPTWSEIEDLRACFGAEP 660
Qy 661 LKKDFSEPLNLD 672
Db 661 LKKDFSEPLNLD 672

RESULT 3
AAV72708
ID AAV72708 standard; protein; 672 AA.
XX AC AAV72708;
XX AC AAV72708;
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (T394A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 394 /note= "Wild type Thr substituted with Ala"
FT EP1077260-A1.
XX PN 21-FEB-2001.
XX PD 13-AUG-1999; 99EP-00115161.
XX PR 13-AUG-1999; 99EP-00115161.
XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PI Nueesch J, Rommelaere J;
XX WP1; 2001-212717/22.
XX DR N-PSDB; AAD02803.
XX PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX PS Claim 6; Page 25-27; 41pp; English.
XX CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX SQ Sequence 672 AA;
Query Match 99.5%; Score 3620; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWQSDMEWETTVDEMTKQVFI FDSLVKKCLFEVLNTKNI FPGDVNWFVQHEW 120

```

```

Db 61 RGAETTWQSDMEWETTVDEMTKQVFI FDSLVKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Qy 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLNLYVWSRWLVTAACNVQLTPAERI KLRREIAEDN 180
Db 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLNLYVWSRWLVTAACNVQLTPAERI KLRREIAEDN 180
Qy 181 EWTLLTYKHQTKDYTKCVLFGNNIAYYFLTKKISTSPRGGYFLSSDSGSKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNNIAYYFLTKKISTSPRGGYFLSSDSGSKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRITQTKKEVAIKTTLKLVLHVKRVTSP 300
Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRITQTKKEVAIKTTLKLVLHVKRVTSP 300
Qy 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTFLARTKTAFDLILKKAETSKLTNFSLPD 360
Db 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTFLARTKTAFDLILKKAETSKLTNFSLPD 360
Qy 361 TRTCRIFAFHGMNYYKVAICCVLANROGKRNVLPHGPASTGKSTIAQAIQAAGVNG 420
Db 361 TRTCRIFAFHGMNYYKVAICCVLANROGKRNVLPHGPASTGKSTIAQAIQAAGVNG 420
Qy 421 CYNAAANVPFNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKGSQKIEPTP 480
Db 421 CYNAAANVPFNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKGSQKIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
Qy 541 NGYQSTWASYCAKWKVPDSENWAEKVPPTPINLLGSARSPETTPKSTPLSQNYALTPL 600
Db 541 NGYQSTWASYCAKWKVPDSENWAEKVPPTPINLLGSARSPETTPKSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETONTGEGSKACODGOLSPTWSEIEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAETONTGEGSKACODGOLSPTWSEIEDLRACFGAEP 660
Qy 661 LKKDFSEPLNLD 672
Db 661 LKKDFSEPLNLD 672

RESULT 4
AAV72710
ID AAV72710 standard; protein; 672 AA.
XX AC AAV72710;
XX AC AAV72710;
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (T463A).
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 463 /note= "Wild type Thr substituted with Ala"
FT EP1077260-A1.
XX PN 21-FEB-2001.
XX PD 13-AUG-1999; 99EP-00115161.
XX PR 13-AUG-1999; 99EP-00115161.
XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

```



QY 241 KGERHLVSKLYTDDMRPETVETTTTAQETKGRGRIQTKKEVAIKTTLKELVHKRVTSPE 300  
 DB 241 KGERHLVSKLYTDDMRPETVETTTTAQETKGRGRIQTKKEVAIKTTLKELVHKRVTSPE 300  
 QY 301 DWMMPQDSYIEMMAQGGNLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360  
 DB 301 DWMMPQDSYIEMMAQGGNLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360  
 QY 361 TRTCRIFAFHGMNKKVCHAIACCVLNRQGGKRNITVLFHGPASTGKSIIAQIAAIVGNVG 420  
 DB 361 TRACRIFAFHGMNKKVCHAIACCVLNRQGGKRNITVLFHGPASTGKSIIAQIAAIVGNVG 420  
 QY 421 CYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQIRIDQKGGSKQIEPTP 480  
 DB 421 CYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQIRIDQKGGSKQIEPTP 480  
 QY 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLVK 540  
 DB 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLVK 540  
 QY 541 NGYQSTMASYCAKWKGVDPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600  
 DB 541 NGYQSTMASYCAKWKGVDPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600  
 QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGAGSKACODGQLSPTWSIEIEDLACFGAEP 660  
 DB 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGAGSKACODGQLSPTWSIEIEDLACFGAEP 660  
 QY 661 LKXDFSEPLNLD 672  
 DB 661 LKXDFSEPLNLD 672

RESULT 6  
 AAP40306 standard; protein; 690 AA.  
 AC AAP40306;  
 DT 24-OCT-2003 (revised)  
 DT 12-JAN-1992 (first entry)  
 XX  
 Sequence encoded by the double-stranded replicative form DNA of porcine parvovirus.  
 KW Protein envelope; immunogen; vaccine; antigen; epitope.  
 XX Porcine parvovirus; NADL-2 virulent strain.  
 PN WO8402847-A.  
 XX  
 PD 02-AUG-1984.  
 XX  
 PF 19-JAN-1984; 84WO-US0000063.  
 XX  
 PR 19-JAN-1983; 83US-00459203.  
 PR 06-JAN-1984; 84US-00567968.  
 XX  
 PA (AMGE-) AMGEN.  
 XX  
 PI Fox GW;  
 XX  
 DR WPI; 1984-201354/32.  
 DR N-PSDB; AAN40252.  
 XX  
 Polypeptide obtd. by recombinant DNA methods - for vaccination against parvovirus infections in man and animals.  
 XX  
 PS Claim 6; Table II, Page 33-49; 80pp; English.  
 XX  
 The inventors claim an immunologically active polypeptide for the development of vaccinal immunity against parvovirus infection. Also

CC claimed are DNA sequences wholly or partly duplicative of defined sequences. The polypeptides are used in vaccines for conferring CC protection against parvovirus infections in man and animals. (Updated on CC 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 690 AA;  
 Query Match 65.2%; Score 2373; DB 1; Length 690;  
 Best Local Similarity 67.7%; Pred. No. 1.7e-222;  
 Matches 448; Conservative 71; Mismatches 127; Indels 16; Gaps 6;  
 QY 2 AGNAYSDVLGATNWLKEKSNQEVFSFVKNNVQNLGKDIGNSYKKEIQEDELKSLQR 61  
 DB 5 AGNTYSEEVLUKATNWLQDNAQKEAFSYVFKTKVNLNGKEIAWNNYNKOTTDAMINLQR 64  
 QY 62 GAETTWQDSQEDMEWETTVDEMTEKKQVFIQFDSLKKCLFEVLNKNTPPGDVNMFVQHEWG 121  
 DB 65 GAETSDQATDMEWESEIDSLTKRQVLIQFDSLKKCLFEGILQKLSPSDCYFQLQHEHG 124  
 QY 122 KDQGWCHVLIGKQDSQAGKWRRLQNLVYWRWLVTACNVQLTPAERIKLREIADNE 181  
 DB 125 QDTGYHCHVLLGGGLQOAMGKWLRLQNLNLSRNLNNQCKVPLTPVERIKLREIADNE 184  
 QY 182 WVTLLTYKHQTKQDYTKCVLFGNMIAYFLTKKJISTSPRDGGYFLSSDSGKTNFLK 241  
 DB 185 WSLLLTYTHKQTKQYTKMTHFGNMIAYFLNKKRKTTE--REHYGLSSDSGFMNFLK 242  
 QY 242 EGRRLVSLKLYTDDMRPETVETTTTAQETKGRGRIQTKKEVAIKTTLKELVHKRVTSPE 301  
 DB 243 EGERHLVSHLFTANKPETVETTTTAQEVPRGRIQTKKEVSIKCTIRDLVNRKCTSI 302  
 QY 302 WMMQPDYSIEMMAQPGGNLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPDT 361  
 DB 303 WSMTPDYSIEMMAQPGGNLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSINT 362  
 QY 362 RTCRIFAFHGMNKKVCHAIACCVL--NRQGGKENTVLFHGPASTGKSIIAQIAAIVGNV 419  
 DB 363 RTCKIFSMNNWYIKVCHAITCVLKQTRKKKYNISCHGPGASTGKSIIAQHIANLVGNV 422  
 QY 420 GCYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQIRIDQKGGSKQIEPT 479  
 DB 423 GCYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQIRIDQKGGSKQIEPT 482  
 QY 480 PVIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLV 539  
 DB 483 PVIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLV 542  
 QY 540 KNGYQSTMASYCAKWKGVDPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALT 599  
 DB 543 KNGYQATMASYMHWNVCNVPDWSKLEPKHSPKHSPTPTDQIS--TSVKTSPTADINYAATP 601  
 QY 600 LASDLE-DLALPWSPTNTPVAGTAETQNTGEGAGSKACODGQL---SPTWSEIEEDLRAC 655  
 DB 602 IQEDLDLALPWSPTNTPVAGTAETQNTGEGAGSKACODGQL---SPTWSEIEEDLRAC 654  
 QY 656 FG 657  
 DB 655 FG 656

RESULT 7  
 AAY58164  
 ID AAY58164 standard; protein; 550 AA.  
 XX  
 AC AAY58164;  
 XX  
 DT 07-MAR-2000 (first entry)  
 XX  
 DE Adeno associated virus AAV5 Rep68 protein.  
 XX  
 KW Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR; promoter; Rep protein; capsid protein; regulation; transcription; replication; chromosomal integration; tissue tropism; cellular receptor;  
 CC

KW gene therapy; neutralising antibody; erythroid progenitor cell;  
KW transduction; cancer; genetic disease; Rep68.  
OS Adeno-associated virus 5.  
XX WO9961601-A2.  
XX PD 02-DEC-1999.  
XX PF 28-MAY-1999; 99WO-US0111958.  
XX XX 28-MAY-1998; 98US-0087029P.  
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Chiorini JA, Kotin RM;  
XX WPI; 2000-062707/05.  
DR N-PSDB; AA249216.  
XX Adeno-associated virus 5 based vectors and particles, useful for gene  
PT therapy.  
XX Claim 17; Page 87-88; 91pp; English.  
XX This sequence represents the Rep68 protein of adeno associated virus type  
CC 5 (AAV5). The invention relates to vectors comprising a pair of AAVS  
CC inverted terminal repeats (ITRs) with a promoter between the ITRs. The  
CC vector may comprise the viral genome, or subregions thereof, including  
CC sequences encoding Rep proteins and capsid proteins, and is encapsidated  
CC in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613),  
CC Rep52 (AAV58168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in  
CC regulation of replication and transcription, in addition to the  
CC production of progeny genomes. Rep68 and Rep78 are also associated with  
CC the stable integration of the viral genome into human chromosomes. The  
CC three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3  
CC (AAV58162) assemble to form an icosahedral capsid, and differ from each  
CC other by the use of alternative splicing and an unusual translation  
CC initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2  
CC capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are  
CC likely to utilise distinct cellular receptors and are serologically  
CC distinct. In a gene therapy application, therefore, AAV5 would allow for  
CC transduction of a patient who already possess neutralising antibodies  
CC either as a result of natural immunological defence or from prior  
CC exposure to AAV2 vectors. The vectors may be useful for transducing  
CC erythroid progenitor cells or cells lacking heparin sulphate  
CC proteoglycans, which is very inefficient with AAV2-based vectors. The  
CC vectors may also be useful for transducing cells with a nucleic acid of  
CC interest in order to produce cell lines that could be used to screen for  
CC agents that interact with the gene product of the nucleic acid of  
CC interest. In addition to transduction of other cell types, transduction  
CC of erythroid cells would be useful or the treatment of cancer and genetic  
CC diseases which can be corrected by bone marrow transplants using matched  
CC donors  
XX  
XX Sequence 550 AA;  
SQ  
Query Match 12.5%; Score 455.5; DB 3; Length 550;  
Best Local Similarity 27.8%; Pred. No. 1e-34;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
QY 67 WQSDMEWETTVDEMTEKQVIFDLSLVKCLFEVLNTKNIFFPGDVNMFVQHEWGKQGW 126  
DB 35 WELPESDLNLTVE--QPQLTVADRIRRVLYE---WNKFSKQESKFFVQFEKGEV-F 88  
QY 127 HCHVLIGKDF-SQAGKWRRLQNVVWSRWLVLTACNVLTAPRIKLRIT-ADENWVT 184  
DB 89 HLTLVETSGISSMVLGRY-----VSQIRAOVLV---KVPFGQIEPQINDWVA 132  
QY 185 LLTYKHQTKDYTKYKFGNMIAYVELTKKISTSPDRDGGYFLSSDSGMK-TNF---- 239  
DB 133 I-----TKYKGGANKVDSGYPAYLLPK-----VQPELOWATNLDYK 173

QY 240 -----LKEGERHLVSKLYTDDMRPETVETTTTAQETKGR-----IOTKKEVAIKTT 287  
DB 174 LAALNLEERKELVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
QY 288 LKELVHKRVTSPEMMQPDSDSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEK 347  
DB 224 VNLVHEGITTSEKOWIQENQESYLSFNSTGNSRQIKAAALDNATKIMSLTKSAVDY---- 279  
QY 348 AETSKLINFSLPD-----TRTCRIFAFHGW--VV-KVCHAICCVLNRQGGKRNVTFLFHP 400  
DB 280 -----LVGSSVPEDISKNRIMQIFEMNGYDPAYAGSILYGC-----QRSFNKENTVWLYGP 331  
QY 401 ASTGKSIITAOIAQAVGNVGCYNAAVNFPPDCTNKNLIWVEAGNFGQOVNOFKAICS 460  
DB 332 ATTGKTNIAEALHTVPPYGCNVNTNENFPDNCVDRKMLIWWEBGKNTNKVYESAKAILG 391  
QY 461 GQTRIDOKGKSGQIEPTPVIMTNTNENITVVRIGCEERPEHTOPIRDRMLNIHLTHLP 520  
DB 392 GSKVRVDQCKSSVQIDSTPVIIVTSNTNMVVDGNSTTFEHOQPLEDRMFKFELTKRLP 451  
QY 521 GDFGLVDKNEMPMICAWLVKNGYQSTWASYCAKWKVKPDWSE-----NMAEPKVP 570  
DB 452 PDFGKITKQEVKDFPFW-----AKYNQVPVTHEFKVPRELACTKGAEKSLK 497  
QY 571 TPI-----NLTGSARSPT--TPKSTPLSONYA-LTFL 600  
DB 498 RPLGVDVTNTSVKSLKRLARLSFVPEPTRSSDVTVDPAPLRPL 539  
RESULT 8  
AAU11409  
ID AAU11409 standard; protein; 550 AA.  
XX AAU11409;  
AC AAU11409;  
XX 26-FEB-2002 (first entry)  
DT Adeno-associated virus 5 (AAV5), Rep68 protein.  
DE Adeno-associated virus 5 (AAV5), Rep68 protein.  
XX Adeno-associated virus 5; AAV5; Rep68; nontropic; neuroprotective;  
KW cytostatic; gene therapy; Parkinson's disease; Alzheimer's disease;  
KW demyelination disease; metabolic disorder; musculoskeletal disease;  
KW cardiovascular disease; cancer; autoimmune disorder; genetic disease;  
KW cystic fibrosis; pseudohypoadosteronism; immobile cilia syndrome;  
KW bronchitis; pneumonia; emphysema; pulmonary oedema;  
KW central nervous system; replication; transcription.  
XX Adeno-associated virus 5.  
OS WO200170276-A2.  
PN 27-SEP-2001.  
XX 22-MAR-2001; 2001WO-US009123.  
XX 22-MAR-2000; 2000US-00533427.  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Chiorini JA, Kotin RM, Davidson B, Zabner J;  
PI WPI; 2002-055104/07.  
XX N-PSDB; AAS17712.  
DR Delivering nucleic acid into cell for treating Parkinson's disease, by the  
PT administering to cell an adeno-associated virus 5 particle comprising the  
PT nucleic acid inserted between a pair of AAV inverted terminal repeats.  
XX Disclosure; Page 126-127; 130pp; English.  
XX The invention describes a novel method of delivering a nucleic acid into  
CC a cell in a subject, comprising administering to the cell an adeno-

CC associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus  
CC which relies on a helper virus for replication, in the absence of which  
CC the AAV5 genome is integrated into a host chromosome in a locus specific  
CC manner. The method provides a way to deliver a nucleic acid to a specific  
CC regions, tissues and cell types of the central nervous system comprising  
CC inserting the nucleic acid between a pair of AAV inverted terminal  
CC repeats or delivering an AAV5 particle containing a vector comprising the  
CC nucleic acid. The method is useful for treating brain disorders such as  
CC demyelination disease, Alzheimer's disease and Parkinson's disease, and  
CC metabolic disorders such as musculoskeletal diseases, cardiovascular  
CC disease, cancer and autoimmune disorders, for treating genetic diseases  
CC such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoparathyroidism,  
CC inotile cilia syndrome, and for treating bronchitis, pneumonia,  
CC emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is  
CC useful for delivering gene that may have a systematic effect like anti-  
CC hypertension drugs, insulin, coagulation factors, antibiotics, growth  
CC factors and hormones. This is the amino acid sequence of the adeno-  
CC associated virus 5 (AAV5) Rep68 protein, one of 4 Rep proteins that  
CC regulate replication and transcription of the AAV5 genome, described in  
CC the method of the invention  
XX  
SQ Sequence 550 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 550;  
Best Local Similarity 27.8%; Pred. No. 1e-34;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
QY 67 WQSDMEWETTVDEMTKKQVFIFDLSLVKKCLFEVLNTKNIFFGDVNVWFVQHEWKGDOGW 126  
DB 35 WELPPSDNLTLVE---QPOLTVADRIRRVFLYE---WNKFSQESKFFVQFEGSEY-F 88  
QY 127 HCHVLIGGKDF-SQAQCKWRRQLNVVMSRWLTACNVQLTPAERIKLREI-AEDNEWVT 184  
DB 89 HLHVLTVETSGISSMVLGRY-----VSQIRALV---KVVFQGIETQINDWA 132  
QY 185 LLYYKHQTKDYTKCVLFGNMTAYELTKKLTSTSPRGGYFLSDSGWK-TNP----- 239  
DB 133 I-----TKVKKGGANKVVDGYIPAYLLPK-----VOPELOWMTNLDYK 173  
QY 240 -----LKEGRHLVSKLYTDMRPETVTTVTAQTKRGR-----IOTKKEVAIKTT 287  
DB 174 LAALNLEERKELVAQFLAESQR-----SQEAASQREPSADPVIKSTSQKYMAL 223  
QY 288 LKELVHKRVTPEDMMMMQPDSEYEMMAQPGENLLKNTLEICTLTARTKTAPDLILEK 347  
DB 224 VNWLVHGHTSEKQWIQENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY---- 279  
QY 348 AETSCLTNFSLPD-----TRTCRIFAFGWN--YV-KVCHAICCVLNRQGGKRVTVLPHGP 400  
DB 280 -----LVGSSVPEDISKNRITWQIFEMNGYDPAYAGSILYGCW---QKSFNKRNTVWLYGP 331  
QY 401 ASTGKSIIAQAIQAVGNCVNAANVFPNDCTKNLWVEAGNFGQOVNQFKAICS 460  
DB 332 ATTGKTNIAIAIAHTVPFYGCVNTNENFPNDCCVDKMLWEEGKNTKNVESAAILG 391  
QY 461 GQTRIDQKQKSGKQIEPTFVIMTNTNITVVRIGCERPEHTQPTIRDRMLNHLTHLTP 520  
DB 392 GSKVRVDQKSSQVIDSTPVIIVTNTNMCVVDGNSSTTFEHQOPLDRMFKFELTKRLP 451  
QY 521 GDFGLVDKNWPMICAWLVKNGYQSTVASTYCAKWKGVDPWSB-----NWAEPKVP 570  
DB 452 PDFGKITQKQEVDFFAW-----AKVNVQVPVTHEFKVPRELATGKAEKSLK 497  
QY 571 TPI-----NILGARSAPT--TPKSTPLSONYA-LTPL 600  
DB 498 RPLGDVNTNTYSKLEKRLAUSFPETPRSSDVTVDPAPLRPL 539

RESULT 9  
AAV58159  
ID AAV58159 standard; protein; 610 AA.  
XX  
AC AAV58159;

XX 07-MAR-2000 (first entry)  
XX Adeno associated virus AAV5 Rep78 protein.  
DE  
XX Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR;  
KW promoter; Rep protein; capsid protein; regulation; transduction;  
KW replication; chromosomal integration; tissue tropism; cellular receptor;  
KW gene therapy; neutralising antibody; erythroid progenitor cell;  
KW transduction; cancer; genetic disease; Rep78.  
XX  
OS Adeno-associated virus 5.  
XX  
XX WO9961601-A2.  
XX  
XX 02-DEC-1999.  
XX  
XX 28-MAY-1999; 99WO-US011958.  
XX  
XX 28-MAY-1998; 98US-0087029P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Chiorini JA, Kotin RM;  
XX WPI; 2000-062707/05.  
DR N-PSDB; AA249211.  
DR  
XX Adeno-associated virus 5 based vectors and particles, useful for gene  
PT therapy.  
XX  
XX Claim 15; Fig 6; 91pp; English.  
XX  
XX This sequence represents the Rep78 protein of adeno associated virus type  
CC 5 (AAV5). The invention relates to vectors comprising a pair of AAV5  
CC inverted terminal repeats (ITRs) with a promoter between the ITRs. The  
CC vector may comprise the viral genome, or subregions thereof, including  
CC sequences encoding Rep proteins and capsid proteins, and is encapsidated  
CC in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613),  
CC Rep52 (AAV59168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in  
CC regulation of replication and transcription, in addition to the  
CC production of progeny genomes. Rep68 and Rep78 are also associated with  
CC the stable integration of the viral genome into human chromosomes. The  
CC three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3  
CC (AAV58162) assemble to form an icosahedral capsid, and differ from each  
CC other by the use of alternative splicing and an unusual translation  
CC initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2  
CC capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are  
CC likely to utilise distinct cellular receptors and are serologically  
CC distinct. In a gene therapy application, therefore, AAV5 would allow for  
CC transduction of a patient who already possess neutralising antibodies  
CC either as a result of natural immunological defence or from prior  
CC exposure to AAV2 vectors. The vectors may be useful for transducing  
CC erythroid progenitor cells or cells lacking heparin sulphate  
CC proteoglycans, which is very inefficient with AAV2-based vectors. The  
CC vectors may also be useful for transducing cells with a nucleic acid of  
CC interest in order to produce cell lines that could be used to screen for  
CC agents that interact with the gene product of the nucleic acid of  
CC interest. In addition to transduction of other cell types, transduction  
CC of erythroid cells would be useful or the treatment of cancer and genetic  
CC diseases which can be corrected by bone marrow transplants using matched  
CC donors  
XX  
SQ Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 3; Length 610;  
Best Local Similarity 27.8%; Pred. No. 1.2e-34;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
QY 67 WQSDMEWETTVDEMTKKQVFIFDLSLVKKCLFEVLNTKNIFFGDVNVWFVQHEWKGDOGW 126  
DB 35 WELPPSDNLTLVE---QPOLTVADRIRRVFLYE---WNKFSQESKFFVQFEGSEY-F 88





AAU11404  
 ID AAU11404 standard; protein; 610 AA.  
 XX AC AAU11404;  
 XX DT 26-FEB-2002 (first entry)  
 XX DE Adeno-associated virus 5 (AAV5), Rep78 protein.  
 XX  
 XX Adeno-associated virus 5; AAV5; Rep78; nootropic; neuroprotective;  
 KW cytostatic; gene therapy; Parkinson's disease; Alzheimer's disease;  
 KW demyelination disease; metabolic disorder; musculoskeletal disease;  
 KW cardiovascular disease; cancer; autoimmune disorder; genetic disease;  
 KW Cystic fibrosis; pseudohypoadosteronism; immature cilia syndrome;  
 KW bronchitis; pneumonia; emphysema; pulmonary oedema;  
 KW central nervous system; replication; transcription.  
 XX OS Adeno-associated virus 5.  
 XX PN WO200170276-A2.  
 XX PD 27-SEP-2001.  
 XX PF 22-MAR-2001; 2001WO-US009123.  
 XX PR 22-MAR-2000; 2000US-00533427.  
 XX PA (IOWA ) UNIV IOWA RES FOUND.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Chiorini JA, Kotin RM, Davidson B, Zabner J;  
 XX NP: 2002-055104/07.  
 XX DR N-PSDB; AAS17712.  
 XX PT Delivering nucleic acid into cell for treating Parkinson's disease, by  
 PT administering to cell an adeno-associated virus 5 particle comprising the  
 PT nucleic acid inserted between a pair of AAV inverted terminal repeats.  
 XX PS Disclosure; Fig 6; 130pp; English.  
 XX  
 XX The invention describes a novel method of delivering a nucleic acid into  
 CC a cell in a subject, comprising administering to the cell an adeno-  
 CC associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus  
 CC which relies on a helper virus for replication, in the absence of which  
 CC the AAV5 genome is integrated into a host chromosome in a locus specific  
 CC manner. The method provides a way to deliver a nucleic acid to a specific  
 CC regions, tissues and cell types of the central nervous system comprising  
 CC inserting the nucleic acid between a pair of AAV inverted terminal  
 CC repeats or delivering an AAV5 particle containing a vector comprising the  
 CC nucleic acid. The method is useful for treating brain disorders such as  
 CC demyelination disease, Alzheimer's disease and Parkinson's disease, and  
 CC metabolic disorders such as musculoskeletal diseases, cardiovascular  
 CC disease, cancer and autoimmune disorders, for treating genetic diseases  
 CC such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoadosteronism,  
 CC immature cilia syndrome, and for treating bronchitis, pneumonia,  
 CC emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is  
 CC useful for delivering gene that may have a systematic effect like anti-  
 CC hypertension drugs, insulin, coagulation factors, antibiotics, growth  
 CC factors and hormones. This is the amino acid sequence of the adeno-  
 CC associated virus 5 (AAV5) Rep78 protein, one of 4 Rep proteins that  
 CC regulate replication and transcription of the AAV5 genome, described in  
 CC the method of the invention  
 XX SQ Sequence 610 AA;  
 Query Match 12.5%; Score 455.5; DB 5; Length 610;  
 Best Local Similarity 27.8%; Pred. No. 1.2e-34;  
 Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
 67 WDQSEDMETTTVDKMTKQVFFDSLVKKCLPEVLNKNIFPCDVMNFWOHMGKQGW 126  
 35 WELPPESDLNLTVE--QPQLTVADRIRRVFLYE---WNKFSQESKFFVQFKGSEY-F 88

127 HCHVLIGGKDF-SQAQKQWRRQLNVVWSRWLTACNVQLTPAERIKLRI-AEDNEWVT 184  
 89 HLHLTVETSGISSMVLGRY-----VSQIRAOQLV---KVVFQIEPQINDWA 132  
 185 LLTYKHQTKDYTKCVLFGNMIAYYFLTKKKISTSPPRGGYFLSSDSGWK-TNF--- 239  
 133 I-----TKVKGKANKVDSGYIPAYLLPK-----VQPELQWANTNLDEYK 173  
 240 -----LKEGERHLVSKLYTDDMRPETVETVTVAQETKGR-----IQTKKEVAIKTT 287  
 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
 288 LKELVHKRVTSPEDDMMMPDSDSIEHMAQPGGNNLLKNTLEICTLTARTKTAFDLILEK 347  
 224 VNWLVHEGHTSEKQWIKENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279  
 348 AETSKLTNFSLPD---TRTCRIFAFHGMN--TV-KVCHAI CCVLRQGGKRNVTLPFHPG 400  
 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAVAGSILYGC---QRSFNKRNVTWLYGP 331  
 401 ASTGKSIIAQIAQAVGNVGCYNAAVNPFNDCTNKNLIWBEAGNGOQVQFKAICS 460  
 332 ATTGKTNI AEIAHTVPFYGCNVNTNENPFNDVCKMLIWEEGKMTNKNVSESAILG 391  
 461 GQTRIDQKGGKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLHP 520  
 392 GSKVRVDQKCKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTPEHQQPLEDRMFKELTKRLP 451  
 521 GDFGLVDKNEPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWAEPKVP 570  
 452 PDFGKITKQEVKDFFAW-----AKNVQPVVTHPEKVPRELACTGKAEKSLK 497  
 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600  
 498 RPLGDTVNTSYKSLKRLSLFVFPETPRSDVTVDPAPLRPL 539

RESULT 12  
 AA22887  
 ID AA22887 standard; protein; 610 AA.  
 XX AC AA22887;  
 XX DT 29-AUG-2003 (revised)  
 XX DT 09-AUG-2002 (first entry)  
 XX DE Adeno-associated virus 5 Rep protein.  
 XX  
 XX Nucleic acid/protein conjugate; NAP; nucleic acid modification; NAM; EAS;  
 KW enzyme attachment sequence; cancer therapy; protein-protein interaction;  
 KW drug discovery; Rep protein; adeno-associated virus; AAV; gene therapy;  
 KW cytostatic; Rep protein.  
 XX Adeno associated virus; 5.  
 XX WO200222826-A2.  
 XX 21-MAR-2002.  
 XX 14-SEP-2001; 2001WO-US028702.  
 XX 14-SEP-2000; 2000US-0232960P.  
 XX (XENC-) XENCOR INC.  
 XX Li M, Melander C, Liu H;  
 XX WPI; 2002-393969/42.  
 XX N-PSDB; AAD36281.  
 XX Library of nucleic acid/protein conjugates, has a fusion of nucleic acid  
 PT modification enzyme and candidate compound, and expression vector having

a fusion of nucleic acids encoding NAM enzyme and the compound.

Disclosure; Fig 21; 96pp; English.

PT The present invention relates to genetic libraries of nucleic acid/  
XX protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic  
PS acid modification (NAM) enzyme (E) and candidate compound), an expression  
XX vector (with a fusion of nucleic acids encoding the enzyme and candidate  
CC protein respectively), an enzyme attachment sequence (EAS; RNA sequence),  
CC where the candidate compound and candidate protein are different and EAS  
CC and the enzyme are covalently linked. The NAP conjugates are useful in  
CC screens to assay binding to target molecules and/or to screen candidate  
CC agents for the ability to modulate the activity of the target molecule.  
CC They are useful in cancer therapy. Sequences of the invention are also  
CC useful to detect protein-protein interaction, in drug discovery, to  
CC discover DNA or nucleic acid binding proteins, using nucleic acids as the  
CC targets and to screen for NAM enzymes with reduced toxicity. NAP conjugates  
CC are also useful in pharmacogenomic studies, for screening bioactive  
CC cells (specifically Rep proteins with reduced toxicity). They are also  
CC agents on surface cells, viruses and microbial organisms. They are also  
CC useful for screening proteins causing phenotypic changes such as  
CC overproduction or inhibition of protein expression, or proteins that  
CC alter attachment, infectivity, etc. of the virus. Sequences of the  
CC invention are also used in gene therapy. The present sequence is adeno-  
CC associated virus (AAV) 5 Rep. (Updated on 29-AUG-2003 to standardise OS  
CC field)

XX Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-34;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

Qy 67 WDQSEDEWETTTVDKTKQVIFDLSLVKKCLPEVLNTKNIIPGDVNVVQVHGWKQDQW 126  
Db 35 WELPPESDLNLTVE--QPOLTVDRIIRRVFLYE--WNKFSKQESKFFVQFEKSEY-F 88  
Qy 127 HCHVLIGKDF-SQAQGWRRQLNVYWSRWLTACNVLTTPAERIKLREI-AEDNEWVT 184  
Db 89 HLHVLVETSGISSMWLGRY-----VSQIRALV---KVVFQIEPQINDWA 132  
Qy 185 LLTVKHKQTKDYTKCVLFNGMIAFYFLTKKKISTSPRDGGVFLSDSGWK-TNF---- 239  
Db 133 I-----TKVKGKANGKVDSCYIPAYLLPK-----VQPELQWAMTNLDEYK 173  
Qy 240 -----LKEGEHLSKLVTDMMRETVETVTTTAQETKGR-----IOTKKEVALKTT 287  
Db 174 LAALNLEERKLVAQFLAESSOR-----SQEAASQREFSADPVKSKTSQKYMAL 223  
Qy 288 LKELVHKVTSPEWMMQPDYSIEMMAQPGENLLKNTLICTLTILARTKTAFLILEK 347  
Db 224 VNLVVEHGTSEKQIENQESYLSFNSTGNSRSGIKAAALDNATKIMSLKSAVDY---- 279  
Qy 348 AETSCLTNFSLPD-----TRTCRIFAFHGNW--YV-KVCHAI CCVLNFGGGRNTVLPHGP 400  
Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC---QRSFNKNTVWLYGP 331  
Qy 401 ASTGKSIITAAIAQVAGNVGCVNAANVPFNDCTNKLIVVEAGNFGQVQVQKALCS 460  
Db 332 ATTKGNTIAEAIAHTVPYFGCVNTNENFPFNDVCDVKMLIWWEGKMTKNVVEAKAILG 391  
Qy 461 GOTIRIDOKGSKOIEPTTPVIMTNNITVVRIGCEERPEHTPIRDRMLNIHLTHLP 520  
Db 392 GSKVRDOKCKSVQIDSTPVLVTSNTNMVVDGNSITFEHQPLEDRMFKFELTKRLP 451  
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKGVDPWSE-----NNAEPKVP 570  
Db 452 PDFGKITTKQEVKDFPAW-----AKNQVFTVTHEFKVPRELACTKGAESLK 497  
Qy 571 TPI-----NILGARSPT--TPKSTPLSQNVA-LTLP 600  
Db 498 RFLGDVNTNTSYKLEKRLSFVPEPRSDVTVDPALRPL 539

RESULT 13

RAE28644  
ID AAE28644 standard; protein; 610 AA.

XX AC AAE28644;

XX 29-AUG-2003 (revised)  
DT 27-DEC-2002 (first entry)

XX Adeno-associated virus 5 Rep protein.

XX Nucleic acid modification enzyme; NAM; enzyme attachment sequence; EAS;  
KW protein design automation; PDA; cancer; protein-protein interaction;  
KW infection; gene therapy; Rep protein.

XX Adeno associated virus; 5.

XX WO200268453-A2.

XX 06-SEP-2002.

XX 19-FEB-2002; 2002WO-US004853.

XX 22-FEB-2001; 2001US-00792629.

XX (XENC-) XENCOR INC.

XX Li M, Dahiyat BI;

XX WPI; 2002-691653/74.

XX N-PSDB; AAD46138.

XX Generating a library of fusion nucleic acids for treating cancer or  
PT infection, or detecting protein-protein interaction, comprises providing  
PT computationally-derived library of candidate protein sequences and  
PT expression vectors.

XX Disclosure; Page 180-182; 246pp; English.

XX The present invention relates to a novel method of generating a library  
CC of fusion nucleic acids. The method involves providing a computationally-  
CC derived library of candidate protein sequences and creating a library of  
CC expression vectors containing a fusion nucleic acid having a sequence  
CC encoding a nucleic acid modification (NAM) enzyme and a sequence encoding  
CC a candidate protein sequence from the library and an enzyme attachment  
CC sequence (EAS) that is recognised by the NAM enzyme. The invention also  
CC relates to the use of a variety of computation methods including protein  
CC design automation (PDA). The method is useful in generating and screening  
CC fusion nucleic acids that may be used in treating cancer or infections,  
CC in detecting protein-protein interactions, discovery of DNA or nucleic  
CC acid binding proteins, protein drug discovery, screening for NAM enzymes  
CC with decreased toxicity to the host cells and NAM enzyme/EAS pairs with  
CC increased affinity or in pharmacogenetic studies. The invention is also  
CC used in gene therapy. The present sequence is Adeno-associated virus 5  
CC Rep protein. This sequence is used to illustrate the method of the  
CC invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-34;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

Qy 67 WDQSEDEWETTTVDKTKQVIFDLSLVKKCLPEVLNTKNIIPGDVNVVQVHGWKQDQW 126  
Db 35 WELPPESDLNLTVE--QPOLTVDRIIRRVFLYE--WNKFSKQESKFFVQFEKSEY-F 88

Qy 127 HCHVLIGKDF-SQAQGWRRQLNVYWSRWLTACNVLTTPAERIKLREI-AEDNEWVT 184  
Db 89 HLHVLVETSGISSMWLGRY-----VSQIRALV---KVVFQIEPQINDWA 132

Qy 185 LLTVKHKQTKDYTKCVLFNGMIAFYFLTKKKISTSPRDGGVFLSDSGWK-TNF---- 239

Db 133 I-----TKVKGANKVDSGYIPAYLLPK-----VQPELOWANTNLDYK 173  
 QY 240 -----LKEGRLHLSKLYTDDMRPETVTTTAQETKGR-----IQTKKEVAIKTT 287  
 Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
 QY 288 LKELVHKRVTSPEWMMQPDSDYIENMAQGGENLLKNTLEICTLTLARTKTAFLDILEK 347  
 Db 224 VNWLVHGITSKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279  
 QY 348 AETSCLTNFSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRVTLFHPG 400  
 Db 280 -----LVGSSVPEDISKNRIWQIFENNGYDPAYAGSILYGC---QRSFNKRVTVWLYGP 331  
 QY 401 ASTGKSIIAQAIQAVGNCYNAANVPFPNDCTNKNLIWVEEAGNFGQGVNQPKAICS 460  
 Db 332 ATTGKTNIABIAHTVPFYGCNVNWTNENFPFNDCTNKNLIWVEEAGNFGQGVNQPKAICS 391  
 QY 461 GQIRIDQKGSQKQIEPTVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLHP 520  
 Db 392 GSKVRVDQKSSQVQIDSTPEVITVTSNTNMCVVVDGNSSTTFHQOPLDRMFKFELTKRLP 451  
 QY 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWABPKVP 570  
 Db 452 PDFGKITKQEVKDFFAW-----AKVQVPTVTHEPKVPRELAGTKGAESKLK 497  
 QY 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600  
 Db 498 RPLGDDVTNTSYKLEKRLARLFSVPETPRSSDVTVDPAFLRPL 539

## RESULT 14

AAE26940  
 ID AAE26940 standard; protein; 610 AA.

AC AAE26940;  
 XX

DT 13-DEC-2002 (first entry)

DE Adeno associated virus 5 Rep protein.

XX Prokaryotic library; candidate protein; nucleic acid modification; NAM;  
 KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;  
 KW enzymology; cosmetic research; toxic; environmental safety assessment;  
 KW nutrient biology; Rep protein.

XX Adeno associated virus.

XX WO200266653-A2.

PN 14-DEC-2001; 2001WO-US049058.

XX 29-AUG-2002.

PF 14-DEC-2001; 2000US-0256163P.

PR 14-DEC-2000; 2000US-0256163P.

XX (XENC-) XENCOR INC.

PA Li M, Liu Y;

XX WPI; 2002-667068/71.

DR N-PSDB; AAD44600.

XX New library of prokaryotic pET-24a expression vectors, host cells or  
 PT nucleic acid/protein conjugates, useful for screening candidate proteins  
 PT and their nucleic acids or modification enzymes for pharmacogenetic  
 PT analysis.

XX Disclosure; Fig 21; 127pp; English.

XX The invention relates to methods and compositions for the construction of  
 CC prokaryotic libraries expressing candidate proteins and the use of these

CC libraries to identify candidate proteins and the nucleic acids encoding  
 CC them. The invention provides a library of prokaryotic pET-24a vectors  
 CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a  
 CC nucleic acid modification (NAM) enzyme or a candidate protein, or a  
 CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or  
 CC the candidate protein, and an enzyme attachment sequence (EAS) recognised  
 CC by the NAM enzyme. The library is used for identifying candidate proteins  
 CC and nucleic acids encoding these proteins, in screening for NAM enzymes  
 CC with decreased toxicity for the host cells, or in identifying novel or  
 CC improved EASs, which may be used for understanding cellular processes or  
 CC any subsequent therapeutic or toxic activities. The nucleic acid/protein  
 CC (NAP) conjugates are useful in diagnostic assays and in research  
 CC including clinical pharmacology, functional genomics, pharmacogenomics,  
 CC agricultural chemicals, environmental safety assessment, chemical sensor,  
 CC nutrient biology, cosmetic research or enzymology. These may also be used  
 CC in vitro screening techniques and in assays with target molecules. The  
 CC present sequence is Adeno associated virus 5 Rep protein used in the  
 CC invention

XX Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 610;  
 Best Local Similarity 27.8%; Pred. No. 1.2e-34;  
 Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSDMEWETTVDENMKQVIFDLSLVKKCLPEVLNTRKNIFPGDVNMFVQHWGKDQGW 126  
 Db 35 WELPPESDLNLTIVE--QPQLTVADRIRRVFLYE--WNKFSQESKFFVQFEKGSSEY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYMSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184

Db 89 HLHTLVETGSSVMILGRY-----VSQIRAOQLV---KVFQGIETPOINDMVA 132

QY 185 LLTYKHQTKDYTKCVLFNGNMIAYFLTKKISTSPRDGGYFLSDSGWK-TNF---- 239

Db 133 I-----TKVKGANKVDSGYIPAYLLPK-----VQPELOWANTNLDYK 173

QY 240 -----LKEGRLHLSKLYTDDMRPETVTTTAQETKGR-----IQTKKEVAIKTT 287

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223

QY 288 LKELVHKRVTSPEWMMQPDSDYIENMAQGGENLLKNTLEICTLTLARTKTAFLDILEK 347

Db 224 VNWLVHGITSKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279

QY 348 AETSCLTNFSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRVTLFHPG 400

Db 280 -----LVGSSVPEDISKNRIWQIFENNGYDPAYAGSILYGC---QRSFNKRVTVWLYGP 331

QY 401 ASTGKSIIAQAIQAVGNCYNAANVPFPNDCTNKNLIWVEEAGNFGQGVNQPKAICS 460

Db 332 ATTGKTNIABIAHTVPFYGCNVNWTNENFPFNDCTNKNLIWVEEAGNFGQGVNQPKAICS 391

QY 461 GQIRIDQKGSQKQIEPTVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLHP 520

Db 392 GSKVRVDQKSSQVQIDSTPEVITVTSNTNMCVVVDGNSSTTFHQOPLDRMFKFELTKRLP 451

QY 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWABPKVP 570

Db 452 PDFGKITKQEVKDFFAW-----AKVQVPTVTHEPKVPRELAGTKGAESKLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600

Db 498 RPLGDDVTNTSYKLEKRLARLFSVPETPRSSDVTVDPAFLRPL 539

## RESULT 15

ABU64865

ID ABU64865 standard; protein; 610 AA.

XX ABU64865;

DT 14-MAY-2003 (first entry)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:50 ; Search time 25 Seconds  
(without alignments)  
1782.628 Million cell updates/sec

Title: US-10-069-056-6  
Perfect score: 3638  
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	437.5	12.0	536	US-09-532-594B-10	Sequence 10, Appl
2	437.5	12.0	623	US-09-532-594B-2	Sequence 2, Appl
3	437.5	12.0	623	US-09-532-594B-11	Sequence 11, Appl
4	433.5	11.9	546	US-09-807-802A-7	Sequence 7, Appl
5	433.5	11.9	623	US-09-807-802A-2	Sequence 2, Appl
6	433.5	11.9	623	US-09-807-802A-5	Sequence 5, Appl
7	430.5	11.8	312	US-09-532-594B-8	Sequence 8, Appl
8	430.5	11.8	399	US-09-532-594B-9	Sequence 9, Appl
9	425.5	11.7	322	US-09-807-802A-11	Sequence 11, Appl
10	425.5	11.7	399	US-09-807-802A-9	Sequence 9, Appl
11	144	4.0	276	US-09-171-461-3	Sequence 3, Appl
12	144	4.0	276	US-09-970-711-3	Sequence 3, Appl
13	122.5	3.4	647	US-09-300-909-17	Sequence 17, Appl
14	113	3.1	589	US-09-107-532A-6086	Sequence 6086, Ap
15	112	3.1	564	US-09-134-001C-2897	Sequence 2897, Ap
16	103	2.8	1503	US-09-677-046A-6	Sequence 6, Appl
17	103	2.8	1509	US-09-677-046A-2	Sequence 2, Appl
18	101	2.8	481	US-09-252-991A-30005	Sequence 30005, A
19	101	2.8	1170	US-09-638-524A-2	Sequence 2, Appl
20	100	2.7	1120	US-09-792-024-95	Sequence 95, Appl
21	99.5	2.7	2409	5180808-2	Patent No. 5180808
22	99	2.7	919	US-09-248-796A-19408	Sequence 19408, A
23	98.5	2.7	556	US-09-248-796A-18763	Sequence 18763, A
24	98.5	2.7	3248	US-08-353-700-1	Sequence 1, Appl
25	98.5	2.7	3248	PCT-US95-16216-1	Sequence 1, Appl
26	98	2.7	605	US-09-107-532A-6077	Sequence 6077, Ap
27	98	2.7	715	US-09-248-796A-20602	Sequence 20602, A

28 97.5 2.7 657 3 US-09-300-909-16 Sequence 16, Appl  
29 97 2.7 291 4 US-09-107-532A-6390 Sequence 6390, Ap  
30 97 2.7 471 4 US-09-134-000C-6228 Sequence 6228, Ap  
31 97 2.7 1115 2 US-08-568-459A-2 Sequence 2, Appl  
32 97 2.7 1115 2 US-08-487-826B-2 Sequence 2, Appl  
33 97 2.7 1115 3 US-09-210-288-2 Sequence 2, Appl  
34 97 2.7 1115 6 5198347-6 Patent No. 5198347  
35 96 2.6 370 4 US-09-583-110-3405 Sequence 3405, Ap  
36 95.5 2.6 441 4 US-09-540-236-2983 Sequence 2983, Ap  
37 95.5 2.6 957 4 US-09-489-039A-11233 Sequence 11233, A  
38 95 2.6 590 4 US-09-134-000C-4755 Sequence 4755, Ap  
39 95 2.6 1307 1 US-08-395-246C-2 Sequence 2, Appl  
40 94.5 2.6 1060 4 US-09-248-796A-16624 Sequence 16624, A  
41 94.5 2.6 2482 1 US-08-328-254-6 Sequence 6, Appl  
42 94.5 2.6 3210 4 US-09-538-092-1154 Sequence 1154, Ap  
43 94 2.6 607 3 US-08-486-099-116 Sequence 116, App  
44 94 2.6 607 3 US-08-484-223B-116 Sequence 116, App  
45 94 2.6 607 3 US-08-919-597-116 Sequence 116, App

#### ALIGNMENTS

RESULT 1  
US-09-532-594B-10  
; Sequence 10, Application US/09532594B  
; Patent No. 6468524  
; GENERAL INFORMATION:  
; APPLICANT: Chorini, John A.  
; APPLICANT: Kotin, Robert M.  
; APPLICANT: Safer, Brian  
; APPLICANT: Davidson, Beverly  
; TITLE OF INVENTION: AAVA VECTOR AND USES THEREOF  
; FILE REFERENCE: 14014.0252U2  
; CURRENT APPLICATION NUMBER: US/09/532,594B  
; CURRENT FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 536  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =  
; OTHER INFORMATION: synthetic construct  
; NAME/KEY: misc feature  
; OTHER INFORMATION: AAVA Rep protein 68  
; US-09-532-594B-10

Query Match 12.0%; Score 437.5; DB 4; Length 536;  
Best Local Similarity 27.7%; Pred. NO. 1.5e-36;  
Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;  
QY 63 AETWQSDMEWETTVDEMTKKQVIFDLSLVKCKLFEVLTKNIFPGDVNVVQHEWKGK 122  
Db 31 AEKEWELPPDSMDLNLIE--QAPLTVAEKLORELFVEWRRVSK--APEALFFVQKEG- 85  
QY 123 DQGWCHCHVLIG-----GKDFSOAQGWRRQLNVYMSRWLTACNVQLTPAERIKL 173  
Db 86 DSYFHLHLIVETGVKSMVGVRSIQKEK-----LVTRIYGVPE----- 126  
QY 174 RETAEDNEWVTLTYKHQ--TKKDYTKCVLFGNMIAYYPLTKKIKISTSPRGGVFLSS 231  
Db 164 ELQWANTMDQYISACLNLAERKRLVAOHL-----THVSOTQEQNKQNPNSD 212  
QY 276 ---IQTKKEVAIKTKLKVHVKVTSPEWMMQPDYSIEMMAQPGGKLNKLTLEICTL 332  
Db 213 APVIRKTSARYMELVGNLVDROITSEKQIQEDQASYSIFNAASNSRSQIKAAALDNASK 272

```
QY 333 TLARTKTAFLDILEKAEKTSKLTNFSLPDTRTCRIFAFHGWNYVVKVCHAI CCVLNRQGGKR 392
   :: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 273 INSLTKTAPDYLVGQNPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGFR 327
   :: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 393 NTVLFHGPASTGKSIIAQIAOAVNGVGCYNAAVNFPPENDCTNKNLIWVEBAGNFGQOV 452
   ||: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 328 NTIWLFGPATTKTNTIAEAIAHVPYGCVNTNENFPENDCVDKMWIWEESGKMTAKV 387
   ||: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 453 NQFKALCSGOTRIDQKGKSGKQIBPTPVMTNENITVVRIGCEBEPHPTOPIRDRMLN 512
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 388 ESAKAILGGSKVRVDQCKSSAQIDPTPVIVTSNTNMCVIDGNSSTTFHQQLQDRMPK 447
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 513 IHLTHLPGDFGLVDKNEWPMICANLVKNGYOSTWASYCAKMG--KVPDWE-NWAEPKV 569
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 448 FELTKLEHDFGKVTQKQEVKDFRWASDHVTEVTFYVRKGGARKRPAPNDADISEPKR 507
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 570 PTPINILGARSPTTPKSTPLSQNYA 596
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 508 ACP-----SVAQPSTSDAEAPV--DYA 527

RESULT 3
US-09-532-594B-2
; Sequence 2, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Davidon, Beverly
; APPLICANT: Davidon, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein (full length)
US-09-532-594B-2

Query Match 12.0%; Score 437.5; DB 4; Length 623;
Best Local Similarity 27.7%; Pred. No. 2e-36;
Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;

QY 63 AETTWDQSEDMETTTVDKTKQVIFDLSLVKKCLFEVLNTKNIIPFGDVNWFVQHEWGK 122
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 31 AEKEWELPPDSMDLNLIE--QAPLTVAEKLOREFLVEWRVSK--APEALFFVQPEKG- 85
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 123 DQGWCHVILG-----GKDFSAQGWRRQLNVVWSRWLTACNVLTQPAERIKL 173
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 86 DSYFHLHILVETGVKSMVGVRSQIKK-----LVTRIYRGVEP----- 126
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 174 RETAEDNEWTLTYKHQ--TKKDYTKCVLFGNMIAYFLTKKKISTSPRGGYFLSS 231
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 127 ----QLPNWFAVTKTRNGAGGNKVDDC-----YIPNLLPK---TQP----- 163
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 232 DSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVTTVTTAETKGR----- 275
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 164 ELQWANTNMDQYISACLNLAERKRLVAQHL-----THVSOTQONKENQNPNSD 212
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 276 ---IQTKKEVAIKTKLKVHKTSPEDWMMQPDYSIEMMAQPGENLLKNTLEICTL 332
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 213 APVIRKTSARYMELVGLVDRGITSEKQWIOEDQASYISFNAASRSRQIKAAALDNASK 272
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 333 TLARTKTAFLDILEKAEKTSKLTNFSLPDTRTCRIFAFHGWNYVVKVCHAI CCVLNRQGGKR 392
   :: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
```

```
Db 273 INSLTKTAPDYLVGQNPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGFR 327
QY 393 NTVLFHGPASTGKSIIAQIAOAVNGVGCYNAAVNFPPENDCTNKNLIWVEBAGNFGQOV 452
   ||: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 328 NTIWLFGPATTKTNTIAEAIAHVPYGCVNTNENFPENDCVDKMWIWEESGKMTAKV 387
   ||: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 453 NQFKALCSGOTRIDQKGKSGKQIBPTPVMTNENITVVRIGCEBEPHPTOPIRDRMLN 512
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 388 ESAKAILGGSKVRVDQCKSSAQIDPTPVIVTSNTNMCVIDGNSSTTFHQQLQDRMPK 447
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 513 IHLTHLPGDFGLVDKNEWPMICANLVKNGYOSTWASYCAKMG--KVPDWE-NWAEPKV 569
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 448 FELTKLEHDFGKVTQKQEVKDFRWASDHVTEVTFYVRKGGARKRPAPNDADISEPKR 507
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 570 PTPINILGARSPTTPKSTPLSQNYA 596
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 508 ACP-----SVAQPSTSDAEAPV--DYA 527

RESULT 3
US-09-532-594B-11
; Sequence 11, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Davidon, Beverly
; APPLICANT: Davidon, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 78
US-09-532-594B-11

Query Match 12.0%; Score 437.5; DB 4; Length 623;
Best Local Similarity 27.7%; Pred. No. 2e-36;
Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;

QY 63 AETTWDQSEDMETTTVDKTKQVIFDLSLVKKCLFEVLNTKNIIPFGDVNWFVQHEWGK 122
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 31 AEKEWELPPDSMDLNLIE--QAPLTVAEKLOREFLVEWRVSK--APEALFFVQPEKG- 85
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 123 DQGWCHVILG-----GKDFSAQGWRRQLNVVWSRWLTACNVLTQPAERIKL 173
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 86 DSYFHLHILVETGVKSMVGVRSQIKK-----LVTRIYRGVEP----- 126
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 174 RETAEDNEWTLTYKHQ--TKKDYTKCVLFGNMIAYFLTKKKISTSPRGGYFLSS 231
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 127 ----QLPNWFAVTKTRNGAGGNKVDDC-----YIPNLLPK---TQP----- 163
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 232 DSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVTTVTTAETKGR----- 275
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 164 ELQWANTNMDQYISACLNLAERKRLVAQHL-----THVSOTQONKENQNPNSD 212
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 276 ---IQTKKEVAIKTKLKVHKTSPEDWMMQPDYSIEMMAQPGENLLKNTLEICTL 332
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 213 APVIRKTSARYMELVGLVDRGITSEKQWIOEDQASYISFNAASRSRQIKAAALDNASK 272
   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 333 TLARTKTAFLDILEKAEKTSKLTNFSLPDTRTCRIFAFHGWNYVVKVCHAI CCVLNRQGGKR 392
   :: ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 273 INSLTKTAPDYLVGQNPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGFR 327
```

Qy	393	NTVLFHGPASTGKSIIIAQAIQAAGNVGCYNAAANVPFPFNDCTNKNLWVEEAGNFQQV	452
Dd	328	NTIWLFGPATTTGTINAEIAHAHVPPFYGCYVNWTNEFPFNDCVDVKVIWWEEGKMTAKVV	387
Qy	453	NQKAICSGQTIRIDOKGKSGKOLEPTPVTMTTNEITTVRIGCESRPEHTQPDIRMLN	512
Dd	388	ESAKAILGGSKVRVDQCKCSAQIDPFPVIVTNTNMCAVIDGNSTTFBHQPLQDRMPK	447
Qy	513	IHLTHLPGFGLVDKNEMPMICAWLVKNGYQSTMASYCAKMG--KVDPHSE--NWAEPKV	569
Dd	448	FELTKRLHFDFGVTKQEVKDDFRWASDHVTEVTHEFFYVRKGGRARKPAPNDADISEPKR	507
Qy	570	PTPINLLGSARSPTTPKSTPLSONYA	596
Dd	508	ACP-----SVAOPSTSDAEPV--DYA	527

## RESULT 4

```

US-09-807-802A-7
; Sequence 7, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 546
; TYPE: PRT
; ORGANISM: AAV-1
; US-09-807-802A-7

```

## RESULT 5

```

US-09-807-802A-2
; Sequence 2, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
; US-09-807-802A-2

```







QY 572 PINLLGSARSPPTT 585  
Db 286 P-----SVADPSTS 294

RESULT 11  
US-09-171-461-3  
; Sequence 3, Application US/09171461  
; Patent No. 6335016  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Adam  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Chioccia, Susanna  
; APPLICANT: Kurzbauer, Robert  
; APPLICANT: Schaffner, Gotthold  
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus  
; FILE REFERENCE: 0652.180000  
; CURRENT APPLICATION NUMBER: US/09/171,461  
; CURRENT FILING DATE: 1999-01-12  
; EARLIER APPLICATION NUMBER: PCT/EP97/01944  
; EARLIER FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: CELO Virus  
; FEATURE:  
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2  
US-09-171-461-3

Query Match 4.0%; Score 144; DB 3; Length 276;  
Best Local Similarity 24.1%; Pred. No. 3.6e-06;  
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;  
  
QY 286 TTLKELVHKRVTSPEWMMQPDY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343  
Db 10 TLVHALIDRGVVSREQWQVDPAQYQFYHRSKQGFK--VRHLRDVIRHMCWSRTLLDY 67  
QY 344 ILEKASTKLTNFSLPD----TRTCRIFAFHGWNTYKVKCHAI CCVLNRQGGKRNITVLFHG 399  
Db 68 MSSASTPS-----PDDVLRNPLYQLLCCNGYNPAVVGTTALIRWAGHQ--SNRNTVWIRG 119  
QY 400 PASTGKSIITAAQIAOAVNGVCYNANVPNDCTNKNLIWVEAGNFGQOVNQFKAIC 459  
Db 120 TPMGAPYLAQAIAYCSPLVGSDVRNKSNNPFGCPDPSLVFWWDGGYVYDCCVGLVKQVF 179  
QY 460 SGTIRIDQK-----KGSQKIEPTVIMTNNITVVRIGCEE--RPEHTQPIRDRMLNTH 514  
Db 180 RGEHVILPPEGLRGNPCSELFRTPVLMYSQADICWTRLRSGLSAEHAVALRDCMYLIR 239  
QY 515 LTHLPGDF---GLVDKNEWPMICAW 537  
Db 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 12  
US-09-970-711-3  
; Sequence 3, Application US/09970711  
; Patent No. 6773709  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Adam  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Chioccia, Susanna  
; APPLICANT: Kurzbauer, Robert  
; APPLICANT: Schaffner, Gotthold  
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus  
; FILE REFERENCE: 0652.1800001  
; CURRENT APPLICATION NUMBER: US/09/970,711  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/171,461  
; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: PCT/EP97/01944  
; PRIOR FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: CELO Virus  
; FEATURE:  
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2  
US-09-970-711-3

Query Match 4.0%; Score 144; DB 4; Length 276;  
Best Local Similarity 24.1%; Pred. No. 3.6e-06;  
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;  
  
QY 286 TTLKELVHKRVTSPEWMMQPDY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343  
Db 10 TLVHALIDRGVVSREQWQVDPAQYQFYHRSKQGFK--VRHLRDVIRHMCWSRTLLDY 67  
QY 344 ILEKASTKLTNFSLPD----TRTCRIFAFHGWNTYKVKCHAI CCVLNRQGGKRNITVLFHG 399  
Db 68 MSSASTPS-----PDDVLRNPLYQLLCCNGYNPAVVGTTALIRWAGHQ--SNRNTVWIRG 119  
QY 400 PASTGKSIITAAQIAOAVNGVCYNANVPNDCTNKNLIWVEAGNFGQOVNQFKAIC 459  
Db 120 TPMGAPYLAQAIAYCSPLVGSDVRNKSNNPFGCPDPSLVFWWDGGYVYDCCVGLVKQVF 179  
QY 460 SGTIRIDQK-----KGSQKIEPTVIMTNNITVVRIGCEE--RPEHTQPIRDRMLNTH 514  
Db 180 RGEHVILPPEGLRGNPCSELFRTPVLMYSQADICWTRLRSGLSAEHAVALRDCMYLIR 239  
QY 515 LTHLPGDF---GLVDKNEWPMICAW 537  
Db 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 13  
US-09-300-909-17  
; Sequence 17, Application US/09300909  
; Patent No. 6306580  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING  
; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR  
; NUMBER OF SEQUENCES: 27  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/300,909  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/083,942  
; FILING DATE: 01-MAY-1998  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 647 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-300-909-17

Query Match 3.4%; Score 122.5; DB 3; Length 647;  
Best Local Similarity 19.7%; Pred. No. 0.0031;  
Matches 110; Conservative 77; Mismatches 173; Indels 197; Gaps 26;  
  
QY 39 GKDIGWNSYKKEIQEDELK-SLQGAETTWQDSE----DMEWETT-VDEMTKQVIFDS 92  
Db 100 GKKGVRNT-RGTLOEISLNVSTQATQTVYSVPDSGYGNVEVTAEEVET----- 149

QY 93 LVKKCLFEVLNTKNIIPGDVNVFVQHEKGDQOQWCHVLIGGKDFSOAQGKWRRLQNVY 152  
Db 150 -----VATNTNGDAEGHGSVREE-----CSSVDSAIDSENQDPK-----185  
QY 153 WSRWLVTACNVQVTPAERIKLREIAEDNEWVTLTKYKHQT-----KKDYTKC 200  
Db 186 -----SPTAQIKLL-LQSNKKAAMLT-QFKETYGLSFTDLVRFKSKTTC 230  
QY 201 VLFGNMIAYYFTKKKISTSPRGGYFLSSDSGWKTNFLKEGERHLVSK--LYTDDMRP 258  
Db 231 T---DWAAIFGVHPTIA-----EGFKTLINKYALYTH---260  
QY 259 ETVEITVTTAETKRGRIOT-----KKEVAIKTTLKELVHKRVTSPEDEMMQPDYSY 310  
Db 261 -----IQSLDTKQGVLLMLIRYTCGRNVTVGKGLSTLLH-----VPESCMLLLEPPKL 309  
QY 311 IEMMA-----OPGENLLKNTLE-----ICTLTLAR---TKTAFDL--ILEKAETSKLT-- 354  
Db 310 RSPVAALYWRIGISNISVVTGTPWIORLTVIQHIGIDSDVPDLSDMVQWAFDNEYTDE 369  
QY 355 -----NFSLPDTRTCTRIFAFHGMN-----YVKVCHAIC-----CVL 385  
Db 370 SDIAFNAMLADCNAAFLKSNCOAKYKDCATWCKHYKRAQKQMSQWIKERCCK 429  
QY 386 NROGG-----KRNVTLPFGHPASTGKSIIAQIAQ 414  
Db 430 CDEGGDRPIVOFLRYOGIEFISFLCALKEFLKGTGPKQNCIVYGANTGKSHFCMSLMH 489  
QY 415 AV-GNVGCYNAANVPFNDCTNKNLWVEAGN--FGQVNOFKALCSGQTRIDOKGK 471  
Db 490 FLOGTVISVNSTSHFWLSEPLADAKLAMLDDATGTCWSYFDNMRNALDGYAISLDRKYK 549  
QY 472 GSKQIEPTVIMTNTEN 488  
Db 550 SLLQMKCPPLITSNTN 566

RESULT 14  
US-09-107-532A-6086  
; Sequence 6086, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6086:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...589  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6086:  
US-09-107-532A-6086

Query Match 3.1%; Score 113; DB 4; Length 589;  
Best Local Similarity 19.9%; Pred: No. 0.026;  
Matches 94; Conservative 60; Mismatches 150; Indels 168; Gaps 23;

QY 162 NVOLTPAERIKUREIAEDNEWVTLT-YKHQTKDYTKCVLFGNMIAYYFTKKKISTS 220  
Db 142 NIKLRNATTKAIRHYLDDNDFLDIETPYLGKSTPEG-----ARDYLVPDRV---187  
QY 221 PRDGGYFLS---SDSGWKNFLKEG--ERHLVSKLYTD-----DMRPE---TVETTVT 266  
Db 188 ---HAGHFYALPQSPQLFKQLLMNAGFDRIYQIVRCFRDEDLGRQPEFTQVDIETFL 244  
QY 267 TAQETKRGRIOTKKEVAIKTTLKELVHKRVTSPEDEMMQPDYSIEMMAQPGGENLLKNT 326  
Db 245 TABE-----IQYTEGLIAKWKVEGVIEVTLFPFM-----TYDEAMARYGSD-----288  
QY 327 LEICTLTLARTKTAFDL-ILEKAETSKLNFSLPDRTRTCRIF--AFHGMNVYKVCCHAIC 383  
Db 289 -----KPDTRFDMELIDLSDTVKEVEF-----KVFQWALENGGVVKALNA---328  
QY 384 VLNRQGGKRNVTLPFGHPASTGKSIIAQIAQAVGNVGCYNAANVPFNDCTNKNLIW---441  
Db 329 -----KGAADRYSRKMDQLGQVYVQFGA-----KGLAWLK 359  
QY 442 VEEAGNFGQVNOFKALCSGQTRIDQKKGSKQIEPTVIM--TTNENITVVRIGCEER 499  
Db 360 VEEDGLKG-PIAKFMGEATEAIK-----ATDAKPGDLLMFGADKSEIVAAALGA---408  
QY 500 PEHTQPIRDRMLNLIHTHLPDGFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKVDP 559  
Db 409 -----IRTR-----LGKELGLIDESKFNL-----428  
QY 560 WSENWAEKVPKVTPIINLLGSARSPTTPKSTPLSONYALTPLASDLSDLELSEP 611  
Db 429 WVTWDPQFEFSEEGRYVSAHHPTMPKE-----EDIPLLAEDP 467

RESULT 15  
US-09-134-001C-2897  
; Sequence 2897, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; ENTEROCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ IDS: 5674  
; SEQ ID NO 2897  
; LENGTH: 564  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:41:01 ; Search time 90.6667 Seconds  
(without alignments)  
2677.792 Million cell updates/sec

Title: US-10-069-056-6  
Perfect score: 3638  
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	455.5	12.5	610	9	US-09-792-630-21
2	455.5	12.5	610	10	US-09-953-351-21
3	455.5	12.5	610	13	US-10-080-376-21
4	455.5	12.5	610	14	US-10-082-671-27
5	455.5	12.5	610	14	US-10-097-100-21
6	455.5	12.5	610	14	US-10-023-208-21
7	455.5	12.5	610	14	US-10-375-192-7
8	452.5	12.4	626	9	US-09-792-630-29
9	452.5	12.4	626	13	US-10-080-376-29
10	452.5	12.4	626	13	US-10-080-376-29
11	452.5	12.4	626	14	US-10-082-671-35
12	452.5	12.4	626	14	US-10-097-100-29
13	452.5	12.4	626	14	US-10-023-208-29

14	452	12.4	627	9	US-09-792-630-25	Sequence 25, Appl
15	452	12.4	627	10	US-09-953-351-25	Sequence 25, Appl
16	452	12.4	627	13	US-10-080-376-25	Sequence 25, Appl
17	452	12.4	627	14	US-10-082-671-31	Sequence 31, Appl
18	452	12.4	627	14	US-10-097-100-25	Sequence 25, Appl
19	452	12.4	627	14	US-10-023-208-25	Sequence 25, Appl
20	445.5	12.2	625	14	US-10-423-704A-3	Sequence 3, Appl
21	437.5	12.0	537	10	US-09-254-747-10	Sequence 10, Appl
22	437.5	12.0	623	9	US-09-792-630-5	Sequence 5, Appl
23	437.5	12.0	623	10	US-09-953-351-5	Sequence 5, Appl
24	437.5	12.0	623	10	US-09-254-747-11	Sequence 11, Appl
25	437.5	12.0	623	10	US-09-254-747-11	Sequence 5, Appl
26	437.5	12.0	623	13	US-10-080-376-5	Sequence 5, Appl
27	437.5	12.0	623	14	US-10-082-671-11	Sequence 11, Appl
28	437.5	12.0	623	14	US-10-097-100-5	Sequence 5, Appl
29	437.5	12.0	623	14	US-10-023-208-5	Sequence 5, Appl
30	437.5	12.0	623	14	US-10-375-192-5	Sequence 5, Appl
31	436.5	12.0	623	9	US-09-792-630-13	Sequence 13, Appl
32	436.5	12.0	623	10	US-09-953-351-13	Sequence 13, Appl
33	436.5	12.0	623	13	US-10-080-376-13	Sequence 13, Appl
34	436.5	12.0	623	14	US-10-082-671-19	Sequence 19, Appl
35	436.5	12.0	623	14	US-10-097-100-13	Sequence 13, Appl
36	436.5	12.0	623	14	US-10-023-208-13	Sequence 13, Appl
37	436.5	12.0	623	14	US-10-375-192-2	Sequence 2, Appl
38	433.5	11.9	546	15	US-10-696-261-7	Sequence 7, Appl
39	433.5	11.9	546	15	US-10-696-282-7	Sequence 7, Appl
40	433.5	11.9	546	15	US-10-696-900-7	Sequence 7, Appl
41	433.5	11.9	623	9	US-09-792-630-11	Sequence 11, Appl
42	433.5	11.9	623	10	US-09-953-351-11	Sequence 11, Appl
43	433.5	11.9	623	13	US-10-080-376-11	Sequence 11, Appl
44	433.5	11.9	623	14	US-10-082-671-17	Sequence 17, Appl
45	433.5	11.9	623	14	US-10-097-100-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-09-792-630-21  
; Sequence 21, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahlyat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RET/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-09-792-630-21

Query Match 12.5%; Score 455.5; DB 9; Length 610;  
Best Local Similarity 27.8%; Pred. No. 1.2e-30;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY	67	WQSEDMEWETTVDENTKQVIFDLSLVKCLPEVLNTKNI	PGDVNWFVQHEWKGQGW	126
DB	35	WELPPESDLNLTAVE--QPQLTVADRIIRVFLYE--	WNKFSQESKFFVQFKGSEY-F	88
QY	127	HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVOLT	PAERIKLREI-AEDNEWVT	184
DB	89	HLHLTVETSGISSWLVGRY-----VSQIRALV---	KVVFQIEQINDWA	132
QY	185	LITYGHQKQTKVCKVLFENMIAYFLTKKISTSP	PRDGGYFLSDSGWK-TNF----	239
DB	133	I-----TKVKGKANKVDSGXIPAYLLPK-----	VQPELQWMTNLDEYK	173
QY	240	-----LKEGERHLVSKLYTDDMRPETVTTVTAQTKGR	-----IQTKKEVAIKTT	287

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
Qy 288 LKELVHKVTSPEDEMMQPDSDYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILLEK 347  
Db 224 VNWLVHEGITSSEKOWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDY---- 279  
Qy 348 AETSKLTNFSLPD-----TRTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400  
Db 280 -----LVGSSVPEDISKRIWIQIFEMNGYDPAYAGSILYGCW---QSFNKNRTVWLYGP 331  
Qy 401 ASTGKSIIAQAIAQAVGNCVNAANVFPFNDCTNKNLIWVEEAGNFQOQVQKKAICS 460  
Db 332 ATTGKTINIAEAIAHTVPFYGCNVNWNENFPFNDCTNKNLIWVEEAGNFQOQVQKKAICS 391  
Qy 461 GOTIRIDOKGKSGKQIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLIP 520  
Db 392 GSKVRVDQCKSSVQIDSTFPIVTSNTNMCVVVDGNSSTTFEHQOPLDRMFKFELTKRLP 451  
Qy 521 GDFGLVDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVP 570  
Db 452 PDFGKITKQEVKDFPAW-----AKVQVPTVTHEFKVPRELAKTGAKESLK 497  
Qy 571 TPI-----NLLGSARSPTT--TPKSTPLSQNYA-LTPL 600  
Db 498 RPLGVDVNTSYKSLERARLSFVPTPRSSDVTVDPAPLRPL 539

## RESULT 2

US-09-953-351-21  
; Sequence 21, Application US/09953351  
; Publication No. US2003036643A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I  
; FILE REFERENCE: A-70814/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/953,351  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232,960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-09-953-351-21

Query Match 12.5%; Score 455.5; DB 10; Length 610;  
Best Local Similarity 27.8%; Pred. No. 1.2e-30;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
Qy 67 WQSEDMEWETTVDGMTKKQVIFDLSLVKCLFEVLNTKNIFFPGDVNMFVQHEWGKDOGW 126  
Db 35 WELPPESDLNLTVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKSEY-F 88  
Qy 127 HCHVLIGKDF--SOAQGWRRQLNVYWSRWLVTCNVQLTPAERIKUREI--AEDNEWT 184  
Db 89 HLHTLVETSGISSMVLGRY-----VSQIRALV---KVVPFGIEPQINDWVA 132  
Qy 185 LLTYKHOTKKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGWK-TNF---- 239  
Db 133 I-----TKVKGANKVVDGSIIPAYLLPK-----VQPELQAWNTLDEYK 173  
Qy 240 -----LKGERHLVSKLYTDDMRPETVETVTTAQTAKRGR-----IOTKKEVAIKTT 287  
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
Qy 288 LKELVHKVTSPEDEMMQPDSDYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILLEK 347  
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
Qy 288 LKELVHKVTSPEDEMMQPDSDYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILLEK 347

Db 224 VNWLVHEGITSSEKOWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDY---- 279  
Qy 348 AETSKLTNFSLPD-----TRTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400  
Db 280 -----LVGSSVPEDISKRIWIQIFEMNGYDPAYAGSILYGCW---QSFNKNRTVWLYGP 331  
Qy 401 ASTGKSIIAQAIAQAVGNCVNAANVFPFNDCTNKNLIWVEEAGNFQOQVQKKAICS 460  
Db 332 ATTGKTINIAEAIAHTVPFYGCNVNWNENFPFNDCTNKNLIWVEEAGNFQOQVQKKAICS 391  
Qy 461 GOTIRIDOKGKSGKQIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLIP 520  
Db 392 GSKVRVDQCKSSVQIDSTFPIVTSNTNMCVVVDGNSSTTFEHQOPLDRMFKFELTKRLP 451  
Qy 521 GDFGLVDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVP 570  
Db 452 PDFGKITKQEVKDFPAW-----AKVQVPTVTHEFKVPRELAKTGAKESLK 497  
Qy 571 TPI-----NLLGSARSPTT--TPKSTPLSQNYA-LTPL 600  
Db 498 RPLGVDVNTSYKSLERARLSFVPTPRSSDVTVDPAPLRPL 539

## RESULT 3

US-10-080-376-21  
; Sequence 21, Application US/10080376  
; Publication No. US20020172968A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahiyat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/080,376  
; CURRENT FILING DATE: 2000-02-19  
; PRIOR APPLICATION NUMBER: US 09/792,630  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-10-080-376-21

Query Match 12.5%; Score 455.5; DB 13; Length 610;  
Best Local Similarity 27.8%; Pred. No. 1.2e-30;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
Qy 67 WQSEDMEWETTVDGMTKKQVIFDLSLVKCLFEVLNTKNIFFPGDVNMFVQHEWGKDOGW 126  
Db 35 WELPPESDLNLTVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKSEY-F 88  
Qy 127 HCHVLIGKDF--SOAQGWRRQLNVYWSRWLVTCNVQLTPAERIKUREI--AEDNEWT 184  
Db 89 HLHTLVETSGISSMVLGRY-----VSQIRALV---KVVPFGIEPQINDWVA 132  
Qy 185 LLTYKHOTKKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGWK-TNF---- 239  
Db 133 I-----TKVKGANKVVDGSIIPAYLLPK-----VQPELQAWNTLDEYK 173  
Qy 240 -----LKGERHLVSKLYTDDMRPETVETVTTAQTAKRGR-----IOTKKEVAIKTT 287  
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
Qy 288 LKELVHKVTSPEDEMMQPDSDYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILLEK 347  
Db 224 VNWLVHEGITSSEKOWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDY---- 279  
Qy 348 AETSKLTNFSLPD-----TRTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400  
Db 280 -----LVGSSVPEDISKRIWIQIFEMNGYDPAYAGSILYGCW---QSFNKNRTVWLYGP 331  
Qy 401 ASTGKSIIAQAIAQAVGNCVNAANVFPFNDCTNKNLIWVEEAGNFQOQVQKKAICS 460



Db 332 ATTGKTNIAEAIHTVFFYGCNVWNTNENFFNDCVDMKLIWEEGKMTNKVVSAAKAILG 391  
Qy 461 GQIRIDQKQKSGQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLHP 520  
Db 392 GSKVRVDQCKSSQVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHOQPLEDRMFKELTKRLP 451  
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKMGKVPDWS-----NWAEPKVP 570  
Db 452 PDFGKITKQEVKDOFFAW-----AKVNOVPVTHEFKVPRELAGTKGAESLK 497  
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600  
Db 498 RPLGVDNTSYKSLEKRLARLSFVPETPRSSDVTVDPAIRPL 539  
RESULT 4  
US-10-082-671-27  
; Sequence 27, Application US/10082671  
; Publication No. US20030049647A1  
; GENERAL INFORMATION:  
; APPLICANT: DAHIYAT, BASSIL  
; APPLICANT: LI, MIN  
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL  
; TITLE OF INVENTION: PROFILES  
; FILE REFERENCE: XEN/001  
; CURRENT APPLICATION NUMBER: US/10/082, 671  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/270, 781  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-10-082-671-27

Query Match 12.5%; Score 455.5; DB 14; Length 610;  
Best Local Similarity 27.8%; Pred. No. 1.2e-30;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
Qy 67 WQSEDMEWETTVDWETKQVFIQVFDLSLVKKCLPEVLNTKNIIPGDVNNVFOVHEWGDQGW 126  
Db 35 WELPPESDLNLTLVE--QPOLTVDADRIRRVFLYE--WNKFSQKESKFFVQPEKSGEY-F 88  
Qy 127 HCHVLIGGKDF-SQAQKWMRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184  
Db 89 HLHVLVETSGISSWVLGRY-----VSQIRAOQLV---KVVFQIEPQINDWA 132  
Qy 185 LLYKHQKQTKDYTKCVLFGNMIAYFLTKKLISTSPRGGYFLSSDSGWK-TNF---- 239  
Db 133 I-----TKVKGKANKVDSGYIPAYLLPK-----VQPELOWAWTNLDEYK 173  
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IOTKEVAIKTT 287  
Db 174 LAALNLEERKRLVAQFLAESSOR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
Qy 288 LKELVHKRVTSPEWMMQPDYSIENMAQPGGENLKNLTLEICTLTLARTKTAFDLILEK 347  
Db 224 VNWLVHGHITSEKQWIQENQESVLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279  
Qy 348 AETSKLTNLSLPD-----TTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRNVTLPFGP 400  
Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAVAGSILYGCW---ORSFNKRNVTWLYGP 331  
Qy 401 ASTGKSIIAQAIAQAGVNGVCYNAANVFPNDCTNKNLIWEEAGNFGQVQNOQKAICS 460  
Db 332 ATTGKTNIAEAIHTVFFYGCNVWNTNENFFNDCVDMKLIWEEGKMTNKVVSAAKAILG 391  
Qy 461 GQIRIDQKQKSGQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLHP 520  
Db 392 GSKVRVDQCKSSQVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHOQPLEDRMFKELTKRLP 451  
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKMGKVPDWS-----NWAEPKVP 570  
Db 452 PDFGKITKQEVKDOFFAW-----AKVNOVPVTHEFKVPRELAGTKGAESLK 497

Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKMGKVPDWS-----NWAEPKVP 570  
Db 452 PDFGKITKQEVKDOFFAW-----AKVNOVPVTHEFKVPRELAGTKGAESLK 497  
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600  
Db 498 RPLGVDNTSYKSLEKRLARLSFVPETPRSSDVTVDPAIRPL 539  
RESULT 5  
US-10-097-100-21  
; Sequence 21, Application US/10097100  
; Publication No. US20030068649A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI  
; FILE REFERENCE: A-70814/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/097,100  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/953,351  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232,960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-10-097-100-21

Query Match 12.5%; Score 455.5; DB 14; Length 610;  
Best Local Similarity 27.8%; Pred. No. 1.2e-30;  
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;  
Qy 67 WQSEDMEWETTVDWETKQVFIQVFDLSLVKKCLPEVLNTKNIIPGDVNNVFOVHEWGDQGW 126  
Db 35 WELPPESDLNLTLVE--QPOLTVDADRIRRVFLYE--WNKFSQKESKFFVQPEKSGEY-F 88  
Qy 127 HCHVLIGGKDF-SQAQKWMRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184  
Db 89 HLHVLVETSGISSWVLGRY-----VSQIRAOQLV---KVVFQIEPQINDWA 132  
Qy 185 LLYKHQKQTKDYTKCVLFGNMIAYFLTKKLISTSPRGGYFLSSDSGWK-TNF---- 239  
Db 133 I-----TKVKGKANKVDSGYIPAYLLPK-----VQPELOWAWTNLDEYK 173  
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IOTKEVAIKTT 287  
Db 174 LAALNLEERKRLVAQFLAESSOR-----SQEAASQREFSADPVIKSKTSQKYMAL 223  
Qy 288 LKELVHKRVTSPEWMMQPDYSIENMAQPGGENLKNLTLEICTLTLARTKTAFDLILEK 347  
Db 224 VNWLVHGHITSEKQWIQENQESVLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279  
Qy 348 AETSKLTNLSLPD-----TTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRNVTLPFGP 400  
Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAVAGSILYGCW---ORSFNKRNVTWLYGP 331  
Qy 401 ASTGKSIIAQAIAQAGVNGVCYNAANVFPNDCTNKNLIWEEAGNFGQVQNOQKAICS 460  
Db 332 ATTGKTNIAEAIHTVFFYGCNVWNTNENFFNDCVDMKLIWEEGKMTNKVVSAAKAILG 391  
Qy 461 GQIRIDQKQKSGQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLHP 520  
Db 392 GSKVRVDQCKSSQVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHOQPLEDRMFKELTKRLP 451  
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKMGKVPDWS-----NWAEPKVP 570  
Db 452 PDFGKITKQEVKDOFFAW-----AKVNOVPVTHEFKVPRELAGTKGAESLK 497

Db 452 PDGKITQEVKDFFAW-----AKNQVPTVTHEFKVPRELAGTKGAESKSLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600

Db 498 RPLGVDVNTSYKSLKRAKRLSFVPETPRSSDVTVDPAPLRPL 539

RESULT 6

US-10-023-208-21

Sequence 21, Application US/10023208

Publication No. US20030124537A1

GENERAL INFORMATION:

APPLICANT: Li, Min

APPLICANT: Liu, Yuan-Ching

FILE OF INVENTION: PROCARYOTIC LIBRARIES AND USES

FILE REFERENCE: A-70174-1/RT/RMS/RMK

CURRENT APPLICATION NUMBER: US/10/023,208

CURRENT FILING DATE: 2001-12-17

PRIOR APPLICATION NUMBER: US 60/256,163

PRIOR FILING DATE: 2000-12-14

NUMBER OF SEQ ID NOS: 63

SOFTWARE: Patentin version 3.1

SEQ ID NO 21

LENGTH: 610

TYPE: PRT

ORGANISM: adeno-associated virus 5

US-10-023-208-21

Query Match 12.5%; Score 455.5; DB 14; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-30;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSEDMEWETTVDKTKQVFIFDLSLVKKCLFEVLNTKNIFPGDVNMFVQHEWGKQDQW 126

Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQPEKGSY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYMSRWLVTACNVOLTPAERIKLREI-AEDNEWT 184

Db 89 HLHVLVETSGISSMWLGRY-----VSQIRAOVL---KVFGQIEPQINDWA 132

QY 185 LUTYKHQKDYTKCVLFGNMIAVYFLTKKISTSPRDGGYFLSSDSGWK-TNF--- 239

Db 133 I-----TKVKGANKVDGSIYPAVLLPK-----VQPELOWAWTNLDEYK 173

QY 240 -----LKEGERHLVSKLYTDDMRPETVETVTTAQTGR-----IQTKEVAIKTT 287

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREBSADPVIKSKTSQKYMAL 223

QY 288 LKELVHKRVTSPEWMMQPDSDYIEMMAOPGGENLLKNLTLEICTLTARTKTAFDLILEK 347

Db 224 VNWLVHGTITSEKQWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY--- 279

QY 348 AETSCLTNFSLPD---TTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400

Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC---QRSFNKNTVWLYGP 331

QY 401 ASTGSIITAAQAAQAVNGVCYNAANVPFPNDCTNKNLIWVEEAGNFQGOVQFKAICS 460

Db 332 ATTGKTINAEIAHTVFFYGCYVNTNENPFNDQVCKMLIWEEGKMTNKKVVEAKAILG 391

QY 461 GQIRIDQKGSQKIEPTVMTTNTNITVVRIGCEERPEHTOPTRDMLNHLTHLPL 520

Db 392 GSKVRVDQKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTFEHOQPLEDRMFELTKRLP 451

QY 521 GDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPPKVP 570

Db 452 PDGKITQEVKDFFAW-----AKNQVPTVTHEFKVPRELAGTKGAESKSLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600

Db 498 RPLGVDVNTSYKSLKRAKRLSFVPETPRSSDVTVDPAPLRPL 539

RESULT 7

US-10-375-192-7

Sequence 7, Application US/10375192

Publication No. US20030224404A1

GENERAL INFORMATION:

APPLICANT: Vega, Manuel

APPLICANT: Dittanti, Lila

TITLE OF INVENTION: HIGH-THROUGHPUT DIRECTED EVOLUTION OF NUCLEIC ACIDS BY RATIONAL

TITLE OF INVENTION: MUTAGENESIS

FILE REFERENCE: 37851-918

CURRENT APPLICATION NUMBER: US/10/375,192

CURRENT FILING DATE: 2003-02-24

PRIOR APPLICATION NUMBER: US 60/360,085

PRIOR FILING DATE: 2002-02-25

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 610

TYPE: PRT

ORGANISM: Adeno-associated virus 5

US-10-375-192-7

Query Match 12.5%; Score 455.5; DB 14; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-30;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSEDMEWETTVDKTKQVFIFDLSLVKKCLFEVLNTKNIFPGDVNMFVQHEWGKQDQW 126

Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQPEKGSY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYMSRWLVTACNVOLTPAERIKLREI-AEDNEWT 184

Db 89 HLHVLVETSGISSMWLGRY-----VSQIRAOVL---KVFGQIEPQINDWA 132

QY 185 LUTYKHQKDYTKCVLFGNMIAVYFLTKKISTSPRDGGYFLSSDSGWK-TNF--- 239

Db 133 I-----TKVKGANKVDGSIYPAVLLPK-----VQPELOWAWTNLDEYK 173

QY 240 -----LKEGERHLVSKLYTDDMRPETVETVTTAQTGR-----IQTKEVAIKTT 287

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREBSADPVIKSKTSQKYMAL 223

QY 288 LKELVHKRVTSPEWMMQPDSDYIEMMAOPGGENLLKNLTLEICTLTARTKTAFDLILEK 347

Db 224 VNWLVHGTITSEKQWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY--- 279

QY 348 AETSCLTNFSLPD---TTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400

Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC---QRSFNKNTVWLYGP 331

QY 401 ASTGSIITAAQAAQAVNGVCYNAANVPFPNDCTNKNLIWVEEAGNFQGOVQFKAICS 460

Db 332 ATTGKTINAEIAHTVFFYGCYVNTNENPFNDQVCKMLIWEEGKMTNKKVVEAKAILG 391

QY 461 GQIRIDQKGSQKIEPTVMTTNTNITVVRIGCEERPEHTOPTRDMLNHLTHLPL 520

Db 392 GSKVRVDQKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTFEHOQPLEDRMFELTKRLP 451

QY 521 GDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPPKVP 570

Db 452 PDGKITQEVKDFFAW-----AKNQVPTVTHEFKVPRELAGTKGAESKSLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600

Db 498 RPLGVDVNTSYKSLKRAKRLSFVPETPRSSDVTVDPAPLRPL 539

RESULT 8

US-09-792-630-29

Sequence 29, Application US/09792630

Patent No. US20020168640A1

GENERAL INFORMATION:

APPLICANT: Li, Min

; APPLICANT: Dahivat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792, 630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-09-792-630-29

Query Match 12.4%; Score 452.5; DB 9; Length 626;  
Best Local Similarity 25.2%; Pred. No. 2.3e-30;  
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;  
QY 75 WETT-----VDEMTKKQVFIFDSLVKKCLPEVLNKNIF-----PGDVNWVQHEWGK 122  
DB 45 WEPTGIWNMEHVNLPMTLADKI-----KNIFQRMNQFNQDETDFFQLEEGS 93  
QY 123 DQGMHCHVLIGGDFSOAQGKWRRLQNVVMSRWLVLTACNVQLTPAERIK---LRBIAED 179  
DB 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129  
QY 180 N-----EWVTLTYKHQKQTKYKCVLFQNMIAYYFLTKKISTSPRRDGGVFLSSDSG 234  
DB 130 KQVKIPDFSIITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWFTNMPLF 179  
QY 235 WKTNFKEGEHLVSKLYTDDMRPEVETVTTTAQETKGRIOQTKKEVAIKTKLKELVHK 294  
DB 180 TAAALCLQKROELDLAFOSEMNNAVQEOQASTAAPL-----ISNRAAKYNSLVLDLIEM 235  
QY 295 RVTSPEDMMQPDYSIEMMAQPGGENLLKNTLEICTLTILARTKTAFLDLILEKAETSKLT 354  
DB 236 GITSEKQMLTENKESYRSQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295  
QY 355 NFSLPDTRCRIPAFHGMVYKCHAI CCVLNRQGGKRNVTLPFGPASTGKSIIAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQYVGSVLCGWVKREFNKNAINWLYGPATTGKTINAEIAH 350  
QY 415 AVGNVCYNAANVPFNDCTNKNLIWBEAGNFGQVNOFKAICSGQTRIDQKKGSK 474  
DB 351 AVPFYGCNVNTNENFPFNDCTNKNLIWBEAGNFGQVNOFKAICSGQTRIDQKKGSK 410  
QY 475 QIEPTVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPCDFGLVDKNE---- 530  
DB 411 CIEPTVITNTDMCMVDGNSSTTMEHRIPLERMFQIVLSHKLGNFGKISKKEVKEF 470  
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568  
DB 471 FKWANDNLVPVVSSEFKVPTNEQTKLTE-----PVPERANEPSPEPKIWAAPTRELEEI 524  
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPSTPNT 617  
DB 525 LRASPELFASVAPLP-----SSDPTSPKRTKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
QY 618 PVAGTAETQNTGAGSKAC-QDQOLSPTWSEIEEDLRACFGAPLKKDFSE 667  
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAPCNELKDDMNE 612

RESULT 9  
US-09-953-351-29  
; Sequence 29, Application US/09953351  
; Publication No. US2003003643A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I  
; FILE REFERENCE: A-70814/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/953, 351  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232, 960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-09-953-351-29

Query Match 12.4%; Score 452.5; DB 10; Length 626;  
Best Local Similarity 25.2%; Pred. No. 2.3e-30;  
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;  
QY 75 WETT-----VDEMTKKQVFIFDSLVKKCLPEVLNKNIF-----PGDVNWVQHEWGK 122  
DB 45 WEPTGIWNMEHVNLPMTLADKI-----KNIFQRMNQFNQDETDFFQLEEGS 93  
QY 123 DQGMHCHVLIGGDFSOAQGKWRRLQNVVMSRWLVLTACNVQLTPAERIK---LRBIAED 179  
DB 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129  
QY 180 N-----EWVTLTYKHQKQTKYKCVLFQNMIAYYFLTKKISTSPRRDGGVFLSSDSG 234  
DB 130 KQVKIPDFSIITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWFTNMPLF 179  
QY 235 WKTNFKEGEHLVSKLYTDDMRPEVETVTTTAQETKGRIOQTKKEVAIKTKLKELVHK 294  
DB 180 TAAALCLQKROELDLAFOSEMNNAVQEOQASTAAPL-----ISNRAAKYNSLVLDLIEM 235  
QY 295 RVTSPEDMMQPDYSIEMMAQPGGENLLKNTLEICTLTILARTKTAFLDLILEKAETSKLT 354  
DB 236 GITSEKQMLTENKESYRSQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295  
QY 355 NFSLPDTRCRIPAFHGMVYKCHAI CCVLNRQGGKRNVTLPFGPASTGKSIIAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQYVGSVLCGWVKREFNKNAINWLYGPATTGKTINAEIAH 350  
QY 415 AVGNVCYNAANVPFNDCTNKNLIWBEAGNFGQVNOFKAICSGQTRIDQKKGSK 474  
DB 351 AVPFYGCNVNTNENFPFNDCTNKNLIWBEAGNFGQVNOFKAICSGQTRIDQKKGSK 410  
QY 475 QIEPTVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPCDFGLVDKNE---- 530  
DB 411 CIEPTVITNTDMCMVDGNSSTTMEHRIPLERMFQIVLSHKLGNFGKISKKEVKEF 470  
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568  
DB 471 FKWANDNLVPVVSSEFKVPTNEQTKLTE-----PVPERANEPSPEPKIWAAPTRELEEI 524  
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPSTPNT 617  
DB 525 LRASPELFASVAPLP-----SSDPTSPKRTKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
QY 618 PVAGTAETQNTGAGSKAC-QDQOLSPTWSEIEEDLRACFGAPLKKDFSE 667  
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAPCNELKDDMNE 612

RESULT 10  
US-10-080-376-29  
; Sequence 29, Application US/10080376  
; Publication No. US20020172968A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahivat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/080, 376  
; CURRENT FILING DATE: 2000-02-19  
; PRIOR APPLICATION NUMBER: US 09/792, 630

```
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-080-376-29

Query Match      12.4%; Score 452.5; DB 13; Length 626;
Best Local Similarity 25.2%; Pred. No. 2.3e-30;
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDENKKQVFIFDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122
DB 45 WEPTGIWNEHVNLPMVTLADKI-----KNIFIQRNQFNQDETDFFFOLEEGS 93
QY 123 DQGWCHVLIGKQFSQAQGWRRQLNVYRWLVTAQNVLTTPAERIK---LREIAD 179
DB 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWVTLTYKHKOTKDYTKCVLFGNMIAYVFLTKKISTSPPRDGGVFLSSDSG 234
DB 130 KQVKIPDWFSTIKTKRGGQNKVTAA-----YILHYLIPKKQ-----PELOWAFTNMPFL 179
QY 235 WKTNFKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVAIKTKLKVHK 294
DB 180 TAAALCLOKQELLDAFOSEMNNAVQEQDASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
QY 295 RVTSPEDEMMQPDSDYIEMMAQPGGENLLKNLTLEICTLTARTKTAFTDLILEKAETSKLT 354
DB 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295
QY 355 NPSLPDTRTCRIFAFHGMVNVKCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIQ 414
DB 296 K-----NRIYQILKLNYPQYVGVLCGWVKREFNKNAINWLYGPATTKGTNIAEIAH 350
QY 415 AVGNVCYNAANVPFNDCNKNLIWVEEAGNFGQOVNQFKAICSGQTIRIDQKGGSK 474
DB 351 AVPFYGVNWTNENFPFNDQVDMLIWEEGKMTNKVESAAILGGSVAVVDQCKGVS 410
QY 475 QIEPTVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNE-----530
DB 411 CIEPTVITSTNDCMIVDGNSTTMEHRIPLERMFQIVLSHKGLENGFKISKKEVKEF 470
QY 531 --W-----PMICAWLVKNGYSTMASYCAKWKGVDPDNGEN-----WAEPK-----568
DB 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTRELEEI 524
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDALEPWPSTPNT 617
DB 525 LRASPELFASVAPLP-----SSPDTSPKPKTKRGEYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETQNTGAGSKAC-QDQGLSPTWSEIEEDLRACFAGBPLKKDFSE 667
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 12
US-10-097-100-29
; Sequence 29, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: LI, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION L
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
```

; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-10-097-100-29

Query Match 12.4%; Score 452.5; DB 14; Length 626;  
Best Local Similarity 25.2%; Pred. No. 2.3e-30;  
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;  
QY 75 WETT-----VDEMTKKQVFIQFDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122  
DB 45 WEPTGIWNEHVNLPMTLADKI-----KNIFQRMNQFNQDDETFDFQLEEGS 93  
QY 123 DQGWCHVLIGGDFSOAQGKWRRLQNVVWSRWLTACNVQLTPAERIK---LREIAED 179  
DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSLRLRDVYEG 129  
QY 180 N-----EWTLLTYKHQKQDYTKCVLFGNMIAYVFLTKKISTSPRRDGGYFLSSDSG 234  
DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSLRLRDVYEG 129  
QY 130 KQVKIPDWFISITTKRGGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179  
DB 235 WKTNFKGGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVAIKTTLKELVHK 294  
DB 180 TAAALCLOKQRELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAAKNSYLNLDVLIEM 235  
QY 295 RVTSPEDMMQPDSDYIENMAOPGGENLLKNTLEICTLTLARTKTAFLDILEKAETSKLT 354  
DB 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDT 295  
QY 355 NPSLPDTRTCRIFAFHGMNYYVVKVCHACCVLNRQGGKRNVTLPFHGPASTGKSIIAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQVYGVSLCGWKEFNKRNALWLYGPATTKTNIAEIAH 350  
QY 415 AVGNVCYNAANVPFNDCTNKLIIWBEAGNFGQOVNQFKAICSGQITRIDQKGGSK 474  
DB 351 AVPFYGVNWTNENFPFNDCTNKLIIWBEAGNFGQOVNQFKAICSGQITRIDQKGGSK 410  
QY 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNE--- 530  
DB 411 CIEPTVITNTDNCMIVDGNSSTTMEHRIPLEERMFQIVLSHLEGNFGKISKEVKEF 470  
QY 531 --W-----PMICAMLVKNGYOSTMASYCAKMGKVPDWSN-----WAEPK----- 568  
DB 471 FKWANDNLVPVSEFVKVPTNEQTKLTE-----PVPERANEPSEPKIWPPTRELEEI 524  
QY 569 -----VPTPINLIGSARSPTTPKSTPLSQNYALTPLASDLLELALPSTPNT 617  
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSEITEEDLRACFGAEPLKDDFSE 667  
DB 577 ERANPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 13  
US-10-023-208-29  
; Sequence 29, Application US/10023208  
; Publication No. US20030124537A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Liu, Yuan-Ching  
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES  
; FILE REFERENCE: A-70174-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/023,208  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/256,163  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 29

; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-10-023-208-29

Query Match 12.4%; Score 452.5; DB 14; Length 626;  
Best Local Similarity 25.2%; Pred. No. 2.3e-30;  
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;  
QY 75 WETT-----VDEMTKKQVFIQFDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122  
DB 45 WEPTGIWNEHVNLPMTLADKI-----KNIFQRMNQFNQDDETFDFQLEEGS 93  
QY 123 DQGWCHVLIGGDFSOAQGKWRRLQNVVWSRWLTACNVQLTPAERIK---LREIAED 179  
DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSLRLRDVYEG 129  
QY 180 N-----EWTLLTYKHQKQDYTKCVLFGNMIAYVFLTKKISTSPRRDGGYFLSSDSG 234  
DB 130 KQVKIPDWFISITTKRGGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179  
QY 235 WKTNFKGGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVAIKTTLKELVHK 294  
DB 180 TAAALCLOKQRELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAAKNSYLNLDVLIEM 235  
QY 295 RVTSPEDMMQPDSDYIENMAOPGGENLLKNTLEICTLTLARTKTAFLDILEKAETSKLT 354  
DB 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDT 295  
QY 355 NPSLPDTRTCRIFAFHGMNYYVVKVCHACCVLNRQGGKRNVTLPFHGPASTGKSIIAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQVYGVSLCGWKEFNKRNALWLYGPATTKTNIAEIAH 350  
QY 415 AVGNVCYNAANVPFNDCTNKLIIWBEAGNFGQOVNQFKAICSGQITRIDQKGGSK 474  
DB 351 AVPFYGVNWTNENFPFNDCTNKLIIWBEAGNFGQOVNQFKAICSGQITRIDQKGGSK 410  
QY 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNE--- 530  
DB 411 CIEPTVITNTDNCMIVDGNSSTTMEHRIPLEERMFQIVLSHLEGNFGKISKEVKEF 470  
QY 531 --W-----PMICAMLVKNGYOSTMASYCAKMGKVPDWSN-----WAEPK----- 568  
DB 471 FKWANDNLVPVSEFVKVPTNEQTKLTE-----PVPERANEPSEPKIWPPTRELEEI 524  
QY 569 -----VPTPINLIGSARSPTTPKSTPLSQNYALTPLASDLLELALPSTPNT 617  
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSEITEEDLRACFGAEPLKDDFSE 667  
DB 577 ERANPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 14  
US-09-792-630-25  
; Sequence 25, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahlyat, Basil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 25  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-09-792-630-25

Query Match 12.4%; Score 452; DB 9; Length 627;  
Best Local Similarity 25.8%; Pred. No. 2.6e-30;  
Matches 168; Conservative 77; Mismatches 266; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVFIPDSLVKCLFEVLTKNIF-----PGDNNWFVQHEWGK 122  
DB 45 WEPTGIWNHVNLPMTLADKI-----KNIFQRMNQFNQDETDFFFQLEEGS 93

QY 123 DQGWCHVLIGGKDFSAQGWRRQLNVYWSRWLTACNVQLTPAERIK---LRIAD 179  
DB 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSG 234  
DB 131 KQVKIPDWSITTKRGGQNKVTAA-----YILHYLIPKQ-----PELOWAFTNMPLF 180

QY 235 WKTNFLKEGRLHVLKLYTDDMRPETVETVTTAQTCKGRIOTKKEVAIKTTLKELVHK 294  
DB 181 TAAALCLQKRELDAFOSENNVAVQEDQASTAAPL-----ISNRAAKNYSNLVDWLIEM 236

QY 295 RVTSPEDDMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354  
DB 237 GIITSEKQMLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 296

QY 355 NFSLPDTRTCRIPAFHGMNYYKVCHAI CCVLNRQGGKRNITVLFHGPASTGKSIIAQAIQ 414  
DB 297 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKNALWLYGPATTKTNIAEIAH 351

QY 415 AVGNVCYNAANVPFNDCNTKNLIWEEAGNFGQOVNQFKAI CSQTIRIDOKGKSK 474  
DB 352 AVPFYGCNVNTNENPFNDCVDKMLIWWEEGKNTKVVESAKAILGGSVAVRDQCKGVS 411

QY 475 QIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNE--- 530  
DB 412 CIEPTPIVITSDMCMIVDGNTTWEHRIPLERMFQIVLSHKLGNFGKISKKEVKEF 471

QY 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568  
DB 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPSEPKIWAAPTREELEEI 525

QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALEPSTPNT 617  
DB 526 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 577

QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSETEEDLRACFGAEPLKDPFSE 667  
DB 578 ERANPFEPQSLGE---NFCNQHG-----WYD-----CAFNCNELKDDMNE 613

RESULT 15  
US-09-953-351-25  
; Sequence 25, Application US/09953351  
; Publication No. US2003036643A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION  
; FILE REFERENCE: A-70814/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/953,351  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232,960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 25  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-09-953-351-25

Query Match 12.4%; Score 452; DB 10; Length 627;

Best Local Similarity 25.8%; Pred. No. 2.6e-30;  
Matches 168; Conservative 77; Mismatches 266; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVFIPDSLVKCLFEVLTKNIF-----PGDNNWFVQHEWGK 122  
DB 45 WEPTGIWNHVNLPMTLADKI-----KNIFQRMNQFNQDETDFFFQLEEGS 93

QY 123 DQGWCHVLIGGKDFSAQGWRRQLNVYWSRWLTACNVQLTPAERIK---LRIAD 179  
DB 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSG 234  
DB 131 KQVKIPDWSITTKRGGQNKVTAA-----YILHYLIPKQ-----PELOWAFTNMPLF 180

QY 235 WKTNFLKEGRLHVLKLYTDDMRPETVETVTTAQTCKGRIOTKKEVAIKTTLKELVHK 294  
DB 181 TAAALCLQKRELDAFOSENNVAVQEDQASTAAPL-----ISNRAAKNYSNLVDWLIEM 236

QY 295 RVTSPEDDMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354  
DB 237 GIITSEKQMLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 296

QY 355 NFSLPDTRTCRIPAFHGMNYYKVCHAI CCVLNRQGGKRNITVLFHGPASTGKSIIAQAIQ 414  
DB 297 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKNALWLYGPATTKTNIAEIAH 351

QY 415 AVGNVCYNAANVPFNDCNTKNLIWEEAGNFGQOVNQFKAI CSQTIRIDOKGKSK 474  
DB 352 AVPFYGCNVNTNENPFNDCVDKMLIWWEEGKNTKVVESAKAILGGSVAVRDQCKGVS 411

QY 475 QIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNE--- 530  
DB 412 CIEPTPIVITSDMCMIVDGNTTWEHRIPLERMFQIVLSHKLGNFGKISKKEVKEF 471

QY 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568  
DB 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPSEPKIWAAPTREELEEI 525

QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALEPSTPNT 617  
DB 526 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 577

QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSETEEDLRACFGAEPLKDPFSE 667  
DB 578 ERANPFEPQSLGE---NFCNQHG-----WYD-----CAFNCNELKDDMNE 613

Search completed: January 22, 2005, 04:03:40  
Job time : 93.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:18 ; Search time 22.3333 Seconds  
(without alignments)  
2895.119 Million cell updates/sec

Title: US-10-069-056-6  
Perfect score: 3638  
Sequence: 1 MAGNAYSDEVLGATNWLKSK.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3625	99.6	672	1 UYPVIM	noncapsid protein
2	3558	97.8	721	1 UYPVIM	noncapsid protein
3	3328	91.5	672	1 UYPVIM	noncapsid protein
4	3324	91.4	668	1 UYPVME	noncapsid protein
5	2649	72.8	668	1 UYPVME	noncapsid protein
6	2646	72.7	668	1 UYPVFP	noncapsid protein
7	2644	72.7	668	1 UYPVCP	noncapsid protein
8	2487	68.4	660	1 UYPVPP	noncapsid protein
9	2478	68.1	662	1 UYPVNA	noncapsid protein
10	1623.5	44.6	392	1 UYPVIF	noncapsid protein
11	550	15.1	641	2 S41439	gene NS-1 protein
12	547.5	15.0	641	2 S41434	gene NS-1 protein
13	543	14.9	620	1 UYPVAP	noncapsid protein
14	533.5	14.7	641	2 S41861	gene NS-1 protein
15	452.5	12.4	626	2 S52209	noncapsid protein
16	424.5	11.7	536	1 UYADIA	noncapsid protein
17	351	9.6	726	1 UYPVSI	noncapsid protein
18	347	9.5	671	1 UYPV19	noncapsid protein
19	168	4.6	490	2 T44050	hypothetical prote
20	165.5	4.5	490	2 J01630	noncapsid protein
21	144	4.0	276	2 S26428	hypothetical 31.5K
22	131.5	3.6	849	1 UYPVAD	noncapsid protein
23	125	3.4	545	2 B44054	orf2 protein - Jun
24	124.5	3.4	614	1 W1ML41	El protein - human
25	122.5	3.4	647	1 W1ML39	El protein - human
26	120.5	3.3	825	2 G96665	protein F22C12.12
27	120.5	3.3	1008	2 T18832	probable RNA helic
28	114.5	3.1	497	2 C70452	transcription regu
29	113.5	3.1	1098	2 B70232	hypothetical prote

ALIGNMENTS

RESULT 1

UYPVIM  
noncapsid protein NS1 - minute virus of mice  
C;Species: minute virus of mice, murine parvovirus  
C;Date: 14-Nov-1983 #sequence\_revision 28-Aug-1985 #text\_change 09-Jul-2004  
C;Accession: A03696  
R;Atell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.  
Nucleic Acids Res. 11, 999-1018, 1983  
A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.  
A;Reference number: A03696; MUID:83143341; PMID:6298737  
A;Accession: A03696  
A;Molecule type: DNA  
A;Residues: 1-672 <AST>  
A;Cross-references: UNIPROT:P03134; EMBL:V01115  
C;Superfamily: parvovirus noncapsid protein  
C;Keywords: noncapsid protein

Query Match 99.6%; Score 3625; DB 1; Length 672;  
Best Local Similarity 99.7%; Pred. No. 1.5e-255;  
Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLGATNWLKSKSQEVSFVKNNVQNGKDIQWNSYKKELODELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKSKSQEVSFVKNNVQNGKDIQWNSYKKELODELKSLQ	60
Qy	61	RGAEITWDQSEDMEWETTVDEMTKKQVFFDLSLVKKCLFEVLTKNIFPGDVNMFVQHEW	120
Db	61	RGAEITWDQSEDMEWETTVDEMTKKQVFFDLSLVKKCLFEVLTKNIFPGDVNMFVQHEW	120
Qy	121	GKQDQGHCHVLIGGKDFSOAQGKWRRLQNLVYWSRWLVACNVQLTPAERIKLRETAEDN	180
Db	121	GKQDQGHCHVLIGGKDFSOAQGKWRRLQNLVYWSRWLVACNVQLTPAERIKLRETAEDN	180
Qy	181	EWVTLTYKHQTKDYTKCVLFGNMIAYVFLTKKXISTSPPRDGGYFLSSDSGKTNFL	240
Db	181	EWVTLTYKHQTKDYTKCVLFGNMIAYVFLTKKXISTSPPRDGGYFLSSDSGKTNFL	240
Qy	241	KEGERHLVSKLYTDDMRPETHVTTAQETKKGRIQTKKEVAIKTKLKLHVKRVTSP	300
Db	241	KEGERHLVSKLYTDDMRPETHVTTAQETKKGRIQTKKEVAIKTKLKLHVKRVTSP	300
Qy	301	DMMMPDSDYIENMAOPGGENLLKNTLEICTLTLATKTAFDILILEKAETSKLTNFSLPD	360
Db	301	DMMMPDSDYIENMAOPGGENLLKNTLEICTLTLATKTAFDILILEKAETSKLTNFSLPD	360
Qy	361	TRTCRIFAFHGMNYYVVKVCHACVCLNRQGGKRNVLPHGPASTGKSIIAQAIAGVNVG	420
Db	361	TRTCRIFAFHGMNYYVVKVCHACVCLNRQGGKRNVLPHGPASTGKSIIAQAIAGVNVG	420
Qy	421	CYNAANVNFNDCTNKNLIWEEAGNFGQVQNFKAICSGQITRIDQKKGSKQIEPTP	480
Db	421	CYNAANVNFNDCTNKNLIWEEAGNFGQVQNFKAICSGQITRIDQKKGSKQIEPTP	480

```
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
|||||
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
|||||
QY 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
|||||
Db 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
|||||
QY 601 ASDLEDLALPWPSTNPTPVAGTAETQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
|||||
Db 601 ASDLEDLALPWPSTNPTPVAGTAETQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
|||||
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672
|||||

RESULT 2
UYPVIM
noncapsid protein NS1 - minute virus of mice (strain MVMi)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A23008; A29510
R:Sahli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A>Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: A23008
A:Molecule type: DNA
A:Residues: 1-721 <SAH>
A:CROSS-references: UNIPROT:Q84363; EMBL:X02481
R:Astell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A>Title: DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and
A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: A29510
A:Molecule type: DNA
A:Residues: 1-645, 1-647-721 <AST>
A:CROSS-references: EMBL:M12032; NID:g332289; PIDN:AAA69566.1; PID:g825477
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 97.8%; Score 3558; DB 1; Length 721;
Best Local Similarity 97.6%; Pred. No. 1.2e-250;
Matches 656; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVQVNGKDIGNWSYKKELQDELKSLQ 60
|||||
Db 50 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKEDVQVNGKDIGNWNYKKELQDELKSLQ 109
|||||
QY 61 RGAETTWQSEDMEWETTVDEMTEKQVFI FDSLVKKCLPEVLTKNI FPGDVNWFVQHEW 120
|||||
Db 110 RGAETTWQSEDMEWETTVDEMTEKQVFI YDLSLVKKCLPEVLTKNIAPADVTFVQHEW 169
|||||
QY 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLQNVYWSRMLVTACNVQLTPAERIKLRETAEN 180
|||||
Db 170 GKQDQGHCHVLIGGKDFSOAQGKWRRLQNVYWSRMLVTACNVQLTPAERIKLRETAEN 229
|||||
QY 181 EWTLLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKNTNLF 240
|||||
Db 230 EWTLLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKNTNLF 289
|||||
QY 241 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
|||||
Db 290 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 349
|||||
QY 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
|||||
Db 350 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 409
|||||
QY 361 TRTCRIFAFHGWNYKVCHAI CCVLRNROGKRNVTLV FHPGASTGKSI IAAIAQAAGNVG 420
|||||
```

```
Db 410 TRTCRIFAFHGWNYKVCHAI CCVLRNROGKRNVTLV FHPGASTGKSI IAAIAQAAGNVG 469
QY 421 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVQVQFKAI CSQTIRIDQKKGSKQIEPTP 480
|||||
Db 470 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVQVQFKAI CSQTIRIDQKKGSKQIEPTP 529
|||||
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
|||||
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 589
|||||
QY 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
|||||
Db 590 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 649
|||||
QY 601 ASDLEDLALPWPSTNPTPVAGTAETQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
|||||
Db 650 ASDLEDLALPWPSTNPTPVAGTAETQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 709
|||||
QY 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721
|||||

RESULT 3
UYPVIM
noncapsid protein NS1 - parvovirus Hi
C:Species: parvovirus Hi
A>Note: host Homo sapiens (man)
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A03695
R:Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A>Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid
A:Reference number: A03695; MUID:83112183; PMID:6823009
A:Accession: A03695
A:Molecule type: DNA
A:Residues: 1-672 <RHO>
A:CROSS-references: UNIPROT:P03133; EMBL:X01457; NID:g60993; PIDN:CAA25689.1; PID:g60994
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.5%; Score 3328; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 5.9e-234;
Matches 612; Conservative 25; Mismatches 35; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVQVNGKDIGNWSYKKELQDELKSLQ 60
|||||
Db 1 MAGNAYSDEVLGATNWLKDKSSQEVFSFVKENVQVNGKDIGNSYRKELQDELKSLQ 60
|||||
QY 61 RGAETTWQSEDMEWETTVDEMTEKQVFI FDSLVKKCLPEVLTKNI FPGDVNWFVQHEW 120
|||||
Db 61 RGAETTWQSEDMEWESA VDDMTKKQVFI FDSLVKKCLPEVLTKNIAPSNVTWVQHEW 120
|||||
QY 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLQNVYWSRMLVTACNVQLTPAERIKLRETAEN 180
|||||
Db 121 GKDPGWHCHVLIGGKDFSQPGQKWRRLQNVYWSRMLVTACNVQLTPAERIKLRETAEN 180
|||||
QY 181 EWTLLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKNTNLF 240
|||||
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYYFLSKKICTSPPRDGGYFLSSDSGKNTNLF 240
|||||
QY 241 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
|||||
Db 241 KEGERHLVSKLYTDEMKEPETVETTVTAQEKRGRIQTRREVSIKTTLKELVHKRVTSPE 300
|||||
QY 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
|||||
Db 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLIANSFMS 360
|||||
QY 361 TRTCRIFAFHGWNYKVCHAI CCVLRNROGKRNVTLV FHPGASTGKSI IAAIAQAAGNVG 420
|||||
Db 361 TRTCRIFAFHGWNYKVCHAI CCVLRNROGKRNVTLV FHPGASTGKSI IAAIAQAAGNVG 420
|||||
```



```
QY 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKSGSKQIEPTP 480
DB 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKSGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNILGARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNILGARSPTTPKSTPLSQNYALTPL 600
QY 601 ASLEDLALPWPSTPNTPVAGTAETONTGEGSKACODGQSLSPTWSEIEBIDLACFCAEP 660
DB 601 ASLEDLALPWPSTPNTPVAGTAETONTGEGSKACODGQSLSPTWSEIEBIDLACFCAEP 660
QY 661 LKKDFSEPLNLD 672
DB 661 LESDFNEELTLD 672

RESULT 4
A44276
noncapsid protein NS1 - parvovirus LuIII
C:Species: parvovirus LuIII
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: A44276
R:Biifoot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A:Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A:Reference number: A44276; MUID:93297126; PMID:8517025
A:Accession: A44276
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-668 <DIF>
A:Cross-references: UNIPROT:P36311; GB:M81888
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.4%; Score 3324; DB 1; Length 668;
Best Local Similarity 91.4%; Pred. No. 1.1e-233;
Matches 614; Conservative 24; Mismatches 30; Indels 4; Gaps 2;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVFNENVOLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVFNEDVQNGKIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWDSQEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
DB 61 RGAETTWDSQEDMEWESSVDLTKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
QY 121 GKQGWCHVLIIGKQFSQAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAEDN 180
DB 121 GKQGWCHVLIIGKQFSQAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAEDQ 180
QY 181 EWTLLTYKHQTKDYKCVLFGNMIAYYFLTKKKISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHQTKDYKCVFGNMIAYYFLTKKKICTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTVAQTKEGRVQTKKEVAIKTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETVTVAQTKEGRVQTKKEVAIKTLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFLDILEKAETSKLTNFSLPD 360
DB 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFLDILEKAETSKLTNFSLPD 360
QY 361 TRTCRIFAFHGMNWKVCHAI CCVLRNQGGRNTVLFHGPASTGKSI IIAQIAQAVGNVG 420
DB 361 TRTCRIFAFHGMNWKVCHAI CCVLRNQGGRNTVLFHGPASTGKSI IIAQIAQAVGNVG 420
QY 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKSGSKQIEPTP 480
```

```
DB 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKSGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNILGARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNILGARSPTTPKSTPLSQNYALTPL 600
QY 601 ASLEDLALPWPSTPNTPVAGTAETONTGEGSKACODGQSLSPTWSEIEBIDLACFCAEP 660
DB 601 ASLEDLALPWPSTPNTPVAGTAETONTGEGSKACODGQSLSPTWSEIEBIDLACFCAEP 660
QY 661 LKKDFSEPLNLD 672
DB 661 LESDFNEELTLD 668

RESULT 5
UYPVME
noncapsid protein NS1 - mink enteritis virus (strain Abashiri)
C:Species: mink enteritis virus, MEV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinaga
J. Gen. Virol. 72, 867-875, 1991
A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: A38350
A:Molecule type: DNA
A:Residues: 1-668 <KAR>
A:Cross-references: UNIPROT:P27438; GB:D00765; NID:g222435; PIDN:BAA00662.1; PID:g222436
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 72.8%; Score 2649; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 1.4e-184;
Matches 491; Conservative 72; Mismatches 103; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVFNENVOLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MSGNQITEEVMEGVNMLKHAENEAFSVFKCDNVQNGKDVHNNYTKFIQNEELTSLI 60
QY 61 RGAETTWDSQ--EDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQH 118
DB 61 RGAETWAMDQTEBEEDMESEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECWFIQH 120
QY 119 EWGKQGWCHVLIIGKQFSQAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAE 178
DB 121 EWGKQGWCHVLIIGKQFSQAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAE 180
QY 179 DNEWVTLTYKHQTKDYKCVLFGNMIAYYFLTKKKISTSPPRDGGYFLSSDSGWKTN 238
DB 181 DSEWVTLTYRHQTKDYKCVVGVNFMNMIAYYFLTKKKI--VHMTKESGYFLSDSGWKN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVETVTVAQTKEGRVQTKKEVAIKTLKELVHKRVTS 298
DB 240 FMKYQDRHVSTLYTSQMKPETVETVTVAQTKEGRVQTKKEVSIKTLRDLVSKRVTS 299
QY 299 PEDMMQPDYSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFLDILEKAETSKLTNFSL 358
DB 300 PEDMMQPDYSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFLDILEKANNTKLTNFDL 359
QY 359 PDTRTCRIFAFHGMNWKVCHAI CCVLRNQGGRNTVLFHGPASTGKSI IIAQIAQAVGN 418
DB 360 ANSRTCQIFRMHGMNWKVCHAIACVLRNQGGRNTVLFHGPASTGKSI IIAQIAQAVGN 419
QY 419 VGCYNAANVNFPPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKSGSKQIEP 478
DB 420 VGCYNAANVNFPPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKSGSKQIEP 479
QY 479 TPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWL 538
```

Db 480 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDKEEPLICAWL 539  
QY 539 VKNGYOSTWASYCAKWKVPDWSNWAEPKVPPIINLLGSARSPFTTPKSTPLSQNYALT 598  
Db 540 VRHGYESTWANYTHHWKVPDWSNWAEPKIQEGV-ISPCKDLETAQAASNPQSDHVL 598  
QY 599 PLASDLEDLALPWPSTPNTPVAGTATONTGEAGSKACODGQLSPTWSIEIEDLRACFGA 658  
Db 599 PLTPDVVDLALPWPSTPNTPIAETA-NQOSNOLG-VTHKDVQASPTWSIEIADLRAIFTS 656  
QY 659 EPLKDFSEPLN 670  
Db 657 EQLEEDFRDLD 668

## RESULT 6

UYPVFP

noncapsid protein NS1 - feline panleukopenia virus (strain 193)

C:Species: feline panleukopenia virus, FPLV

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004

C:Accession: A36608

R:Martin, J.C.; Davidson, B.E.; Studdert, M.J.

J. Gen. Virol. 71, 2747-2753, 1990

A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus

A:Reference number: A36608; MUID:91073139; PMID:2174965

A:Accession: A36608

A:Molecule type: DNA

A:Residues: 1-668 &lt;VAR&gt;

A:Cross-references: UNIPROT:P24842; GB:X55115; NID:G60863; PIDN:CAA38910.1; PID:G60864

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 72.7%; Score 2646; DB 1; Length 668;

Best Local Similarity 73.1%; Pred. No. 2.3e-184;

Matches 491; Conservative 72; Mismatches 103; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNWLKESNOEVFSFVKFNENVLQNGKDIGNSYKKELQBELKSLQ 60

Db 1 MSGNQYTEEVMEGVNWLKHAENEAFSFVKCDNVQLNGKDVWRNNTKPIQNEELTSLI 60

QY 61 RGAETTTWDS--EDMEWETTVDEMTEKKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNMFVQH 118

Db 61 RGAQTAMDTQTEEBEEMDWESEVDSLAKKQVOTFDALIKKCLFEVFSKNIEPNECVWFIOH 120

QY 119 EWKDGQGWCHVLIGKDPQAOQKWRRLQNVVWSRWLVACNVOLTPAERIKLREIAE 178

Db 121 EWKDGQGWCHVLHSHKNLQQAATGKWLRRQNNYWSRWLVTLCSVNLTPTEKIKLREIAE 180

QY 179 DNEWVTLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPRPGGYFLSSDSGWKTN 238

Db 181 DSEWVILTYRHQTKDYKDYVVMVHFGNMIAYYFLTKKII-VHMTKESGYFLSTDGSKFN 239

QY 239 FLKEGERHLVSKLYTDDMRPETVETVTTAQTGKRIQTKKEVAIKTTLKELVHGRVTS 298

Db 240 FMKYQDRHVTSTLYTEQMKPETVETVTTAQTGKRIQTKKEVSIKTLRDLVSKRVTS 299

QY 299 PEDWMMQDPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAEISKLTNFSL 358

Db 300 PEDWMMQDPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKADNTKLTNFDL 359

QY 359 PDTRTCRIFAFHGWNVYKVCCHAICCVLNROGKRNVTLPFGPASTGKSIQAIAQAVGN 418

Db 360 ANSRTQCFPMHGNWIKVCHAIACVLRQGGKNTVLPFGPASTGKSIQAIAQAVGN 419

QY 419 VGCYNAANVPFPNDCTNKNLIWVEEAGNFGQVNOFKAICSGQTTIRIDQKGGSKQIEP 478

Db 420 VGCYNAANVPFPNDCTNKNLIWEEAGNFGQVNOFKAICSGQTTIRIDQKGGSKQIEP 479

QY 479 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDKNEPMLICAWL 538

Db 480 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDKEEPLICAWL 539

QY 539 VKNGYOSTWASYCAKWKVPDWSNWAEPKVPPIINLLGSARSPFTTPKSTPLSQNYALT 598  
Db 540 VRHGYESTWANYTHHWKVPDWSNWAEPKIQEGV-ISPCKDLETAQAASNPQSDHVL 598  
QY 599 PLASDLEDLALPWPSTPNTPVAGTATONTGEAGSKACODGQLSPTWSIEIEDLRACFGA 658  
Db 599 PLTPDVVDLALPWPSTPNTPIAETA-NQOSNOLG-VTHKDVQASPTWSIEIADLRAIFTS 656  
QY 659 EPLKDFSEPLN 670  
Db 657 EQLEEDFRDLD 668

## RESULT 7

UYPVCP

noncapsid protein NS1 - canine parvovirus (strain N)

C:Species: canine parvovirus, CPV

C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004

C:Accession: A29962

R:Reed, A.P.; Jones, E.V.; Miller, T.J.

J. Virol. 62, 266-276, 1988

A:Title: Nucleotide sequence and genome organization of canine parvovirus.

A:Reference number: A29962; MUID:88062992; PMID:2824850

A:Accession: A29962

A:Molecule type: DNA

A:Residues: 1-668 &lt;REE&gt;

A:Cross-references: UNIPROT:P12929; EMBL:M12926; NID:G333438; PIDN:AAA67459.1; PID:G33344

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 72.7%; Score 2644; DB 1; Length 668;

Best Local Similarity 73.1%; Pred. No. 3.3e-184;

Matches 491; Conservative 72; Mismatches 103; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNWLKESNOEVFSFVKFNENVLQNGKDIGNSYKKELQBELKSLQ 60

Db 1 MSGNQYTEEVMEGVNWLKHAENEAFSFVKCDNVQLNGKDVWRNNTKPIQNEELTSLI 60

QY 61 RGAETTTWDS--EDMEWETTVDEMTEKKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNMFVQH 118

Db 61 RGAQTAMDTQTEEBEEMDWESEVDSLAKKQVOTFDALIKKCLFEVFSKNIEPNECVWFIOH 120

QY 119 EWKDGQGWCHVLIGKDPQAOQKWRRLQNVVWSRWLVACNVOLTPAERIKLREIAE 178

Db 121 EWKDGQGWCHVLHSHKNLQQAATGKWLRRQNNYWSRWLVTLCSVNLTPTEKIKLREIAE 180

QY 179 DNEWVTLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPRPGGYFLSSDSGWKTN 238

Db 181 DSEWVILTYRHQTKDYKDYVVMVHFGNMIAYYFLTKKII-VHMTKESGYFLSTDGSKFN 239

QY 239 FLKEGERHLVSKLYTDDMRPETVETVTTAQTGKRIQTKKEVAIKTTLKELVHGRVTS 298

Db 240 FMKYQDRQIVSTLYTEQMKPETVETVTTAQTGKRIQTKKEVSIKTLRDLVSKRVTS 299

QY 299 PEDWMMQDPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAEISKLTNFSL 358

Db 300 PEDWMMQDPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKADNTKLTNFDL 359

QY 359 PDTRTCRIFAFHGWNVYKVCCHAICCVLNROGKRNVTLPFGPASTGKSIQAIAQAVGN 418

Db 360 ANSRTQCFPMHGNWIKVCHAIACVLRQGGKNTVLPFGPASTGKSIQAIAQAVGN 419

QY 419 VGCYNAANVPFPNDCTNKNLIWVEEAGNFGQVNOFKAICSGQTTIRIDQKGGSKQIEP 478

Db 420 VGCYNAANVPFPNDCTNKNLIWEEAGNFGQVNOFKAICSGQTTIRIDQKGGSKQIEP 479

QY 479 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDKNEPMLICAWL 538

Db 480 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDKEEPLICAWL 539

QY 539 VKNGYOSTWASYCAKWKVPDWSNWAEPKVPPIINLLGSARSPFTTPKSTPLSQNYALT 598

Db 540 VRHGYESTWANYTHHWKVPDWSNWAEPKIQEGINSFG-CKDLETAQAASNPQSDQVLT 598



Db 421 YNAANVNPFFDCTNKNLWTEEAGNSQVNOQKAIKSGQTIKIDQKGSQKIEPTPV 480  
 Qy 482 IMTINENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVKN 541  
 Db 481 IMTINEDITKVRIGCEERPEHTQPIRDRMLNINTRKLPDGLGTEWELICAWLVKK 540  
 Qy 542 GYQSTMASYCAKGVKPDWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYALTPLA 601  
 Db 541 GYQATMASYMHGWNVDWSEKWEKPKQTPINTDQSIS-TSVKTSPADNNVAAPIQ 599  
 Qy 602 SDLE-DLALPWSPTNPVAGTASTQNTGEAGSKACODGQL---SPWSEIEEDLRCFG 657  
 Db 600 EDLDLALPWSPTNPFTTNLHLTTP-----PDSAIRTPSPWSEIETDIRACFG 652

## RESULT 10

UYFVLF

noncapsid protein NS1 - feline panleukopenia virus (fragment)

C:Species: feline panleukopenia virus, FPLV

C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004

C:Accession: A03697

R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.

J. Virol. 55, 574-587, 1985

A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parv

A:Reference number: A03697; MUID:85265017; PMID:2991581

A:Accession: A03697

A:Molecule type: DNA

A:Residues: 1-392 &lt;CAR&gt;

A:Cross-references: UNIPROT:P06431; EMBL:M10824; NID:g333474; PIDN:AAA47160.1; PID:g3334

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 44.6%; Score 1623.5; DB 1; Length 392;  
 Best Local Similarity 77.5%; Pred. No. 2.5e-110;  
 Matches 306; Conservative 34; Mismatches 52; Indels 3; Gaps 3;

Qy 276 IQTKKEVAIKTKLKHVKRVTSPEDMMQPDYSIEMMAQPGENLLKNTLLEICTLTLA 335  
 Db 1 IQTKKEVSIKTLRLVSKRVTSPEDMMQPDYSIEMMAQPGENLLKNTLLEICTLTLA 60  
 Qy 336 RTKTAPELLEKAEKSKLTNPSLPDRTCRIFAFHGWNVYKVCIAICCVLNROGGKNTV 395  
 Db 61 RTKTAPELLEKANNKTLNFDLANSRTQIFRMHGWNWIKVCHALACVNLROGGKNTV 120  
 Qy 396 LFHGPASTGKSIIAQAIAQAVGNVGCYNAANVPFNDCTNKLIIWVEAGNFGQVNOF 455  
 Db 121 LFHGPASTGKSIIAQAIAQAVGNVGCYNAANVPFNDCTNKLIIWVEAGNFGQVNOF 180  
 Qy 456 KAICSGQTIKIDQKGSQKIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHL 515  
 Db 181 KAICSGQTIKIDQKGSQKIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIKL 240  
 Qy 516 THHLPGDFGLVDKNWPMICAWLVKQYQSTMASYCAKGVKPDWSENWAEKVPPTPINL 575  
 Db 241 VCKLPDGLVDKEEPMPLICAWLVKHYESTMANYTHHWGKLPWEDENWAEKPIQGVNS 300  
 Qy 576 LGSARSPFTTPKSTPLSQNYALTPLASDLELALPWSPTNPVAGTASTQNTGEAGSKA 635  
 Db 301 PG-CKOLETQAASNPQSDHVLPLTPDVEDLALPWSPTDPIETA-NQGSNQLG-VT 357  
 Qy 636 CQDQGLSPWSEIEEDLRCFGAEPLKQFSEPLN 670  
 Db 358 HKDVQASPTWSEIEADLRAIFTSEQLEEDPRDDL 392

## RESULT 11

S41439

gene NS-1 protein - Aleutian mink disease virus

C:Species: Aleutian mink disease virus

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S41439

R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.

submitted to the EMBL Data Library, January 1994

A:Reference: Sequence comparison of the non-structural genes of four different types of

A:Accession: S41439

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-641 &lt;GOT&gt;

A:Cross-references: UNIPROT:Q65017; EMBL:Z29576

C:Genetics:

A:Introns: 586/1

C:Superfamily: parvovirus noncapsid protein

Query Match 15.1%; Score 550; DB 2; Length 641;  
 Best Local Similarity 25.8%; Pred. No. 6.3e-32;  
 Matches 178; Conservative 108; Mismatches 262; Indels 142; Gaps 25;

Qy 18 KEKSNQEVFSFVKENVOLNGKIDGWSYKKELEQDE-----LKSQRGAETWTD 68  
 Db 8 EORRLQDLYVLKKEIN--DGEVAVLFOQTYTDDKONKPKATPPLTATSSDLRLAFD 64  
 Qy 69 QSEDMEWETTVDWTK-----KQVTFPDSLVKKCLEFVLNTKNIFFDGDVNNWFO 117  
 Db 65 SIEE-NLTASNQOLTNNGINFCKLTLGLTLLLDLKHVSHRWD--NNK-----VNLIWQ 115  
 Qy 118 HEMGKQDQWHCHVLIG-----GKDFSOAQGWRRQLN-----VYWSRWLVLTACNVQ- 164  
 Db 116 IEKGTQOQPHIHCCILGYFDRNEDPKDVOKSLG-WFMKRLNKDLALIYNH-----HCDIQD 170  
 Qy 165 -LTPAERIKLRIADENWVTLTKHKQTKDYKCVLFGNMIAYFYFLTKKISTSPPR 223  
 Db 171 IKDPEDRAKNLKVWVEDGTPKPYKFNQTKQDYKNPKVHLRDYTFIYLFNKKDKINTDSM- 229  
 Qy 224 DGYFLSSDSGWKTNFLKEGRHLYSKLYTDDMRPETVETTV-----TTAQ 269  
 Db 230 -DGYFAAGGIVDN-LTNKERKTLRKMLDQSSIDMDANIDWEDGQDAPKVTDQDQA 287  
 Qy 270 ETKRG-----RIQTKKEVA-----IKTKELVHKRVTSPEWMMQ 307  
 Db 288 TTKTGTSLIWKSCATKVTSKKEVANPVQPSKKLYSAQNTLDALFNVGCTBEDMIKQS 347  
 Qy 308 DSYIEMMAQPGENLLKNTLEICTLARTKTAFLILE--KAETSCLTNFSLPDRICR 365  
 Db 348 DKYLSLEPNFPQKINTLLHNNQVKTSTMTAFDCIIFNEEEDDKPLATIKDM---- 403  
 Qy 366 IFAPFGWNVKVCIAICCVLNROGKNTVLFHGPASTGKSIIAQAIAQAVGNVCYNA 425  
 Db 404 -----GLAEQYLKVLCTILTKQGRGCIWYFGGCTGKTLASLICKATVYGMVTS 458  
 Qy 426 NVNPFNDCTNKLIIWVEAGNFGQVNOFKAICSGQTIKIDQKGSQKIEPTPVIMT 485  
 Db 459 NPNFPWTDGNNRIIWAEECGNLGNWVEDFKAITGGGVKVDTKNKPQSIKGC-VIVTS 517  
 Qy 486 NENITVVRIGCEERPEHTQPIRDRMLNIHL-----THHLPGDFGLVDKNWPMICAWL 538  
 Db 518 NTNITKVTGVCEVTHAEPKQMKIRCMKTIINPKTKITFG-----MLKRWL 566  
 Qy 539 VKNGYQSTMASYCAKGVKVP-DWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYAL 597  
 Db 567 -----NTWDRQPIQLSHEMPELVLET-----TGNSSATTAATKNTGNSQ----- 605  
 Qy 598 TPLASDLELALPWSPTNPVAGTASTQNT 627  
 Db 606 PTTAKSAESVNTNCDTPKRGASSVPPKQH 635

## RESULT 12

S41434

gene NS-1 protein - Aleutian mink disease virus

C:Species: Aleutian mink disease virus

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S41434

R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.

submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus  
A;Reference number: S41434  
A;Accession: S41434  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-641 <GOT>  
A;Cross-references: UNIPROT:Q65020; EMBL:Z29577  
C;Genetics:  
A;Introns: 586/1  
C;Superfamily: parvovirus noncapsid protein

Query Match 15.0%; Score 547.5; DB 2; Length 641;  
Best Local Similarity 25.3%; Pred. No. 9.6e-32;  
Matches 183; Conservative 109; Mismatches 260; Indels 171; Gaps 26;

Qy 18 KEKSNQEVFVFVKENNVQLNGDKIGWNSYKKELOEDE-----LKSQRGAETWTD 68  
Db 8 EQRRLQELFE-KFKTE--VADGEGLAWLFQKQVTDKDNKPTKATPLRTTSSDLRLAFD 64  
Qy 69 QSEDMEWETTVDKTKQVFIQSLVKKCLFEVLNTKNI-----PGDVNMFVQHEWKGKD 123  
Db 65 SIEE-TLKTSNQCLTNNDINFCFLTLGKTL--VLLDKHVRSHRDANKVNFVQVEKGKT 121  
Qy 124 QGWHCHVLIG-----GKDFSOAQGWRRQLN-----VYMSRWLVYACNVOLTPAERIK 172  
Db 122 QPHIHCCLGYFKOEDSKDVQKSLG-WFKLKNKOLAVIWSNHCDIQIGQSDGRADN 180  
Qy 173 LBREAEDNEWVTLTYKHKOTKDYTKCVLFGNMIAYFLTKKISTSPRDRG--GYFLS 230  
Db 181 LKIWIEDGP-TKPKYFNKQKDYKQVFNKPTNLRYTLIYLFNKDKIT-----QEGMDGYAS 235  
Qy 231 SDSGWNTNFKGGERHLVSKLYTDDM-----RPETVETVTTAQETKRG 274  
Db 236 GNGGIIDN-LTNKERRALKRWYLDQSQGILDEDWEDSQSAPKVTDDQSDSATSKTGS 294  
Qy 275 -----RIQTKKEVA-----KTLKELVHKRVTSPEDEMMQSDSIEMM 314  
Db 295 LVWKSATKVTSSKEVAIPVKQPSKQWTSQNTLDDLYMFCGCTPEDMLKMSDRYLEMS 354  
Qy 315 AOPGGENLLKNTLEICTLTARTKTAPDLILE--KAETSKLTNFSLPDTRTCRIFAFHGW 372  
Db 355 LEPNGAKINTLLHMNQVRTSSNLTAFECIIKNEEBEDKPLDTIKOM-----GL 405  
Qy 373 NYVKVCHAICCVLNRQGGKNTVLFHGPASTGKSIITAOIAQAVNGVGVNAAVNPFPFN 432  
Db 406 NEGLKKVLCTILTQSGKRGCVWFYFGPGTGKTLASLICKAVVNYGMVTTSNPNFPWT 465  
Qy 433 DCTNKNLIWVEBAGNFQOVNQFKAICSQGTIRIDQKGSQKQIEPTPTVIMTNNITVV 492  
Db 466 DCGNRNIIWABECGNIGNYVEDFKAITGGDVKVDTKNKPQSIKGS-VIVTSNTNITKV 524  
Qy 493 RIGCEERPEHTQPIRDRMLNIHL-----THLPGDFGLVDKNEWPMICAWLVKNGYQS 545  
Db 525 TVCGVEINVAEPLKQRMVKIRCKMKNPKTKITPG-----MLKWTI----- 566  
Qy 546 TMASYCAKQKVPDWSNNAEPKVPPTINLLGSARSPTTPKSTPLSQNVALTPLASDLE 605  
Db 567 -----STWDRIIP-----IKLSHEMP 581  
Qy 606 DLALPWSSTNTVPAGTAETONTGEAGSKACQDGLSPTWSEIEEDLRACFGAEPKDF 665  
Db 582 ELYLET-SGPNSS-SATTATKST-----GNLQPTTAETAESVNTANCDDTPKRGAS 628  
Qy 666 SEP 668  
Db 629 SVP 631

RESULT 13  
UYVPAP

noncapsid protein NS1 - Aleutian mink disease virus (strain ADV-G)  
N;Alternate names: left-ORF protein  
C;Species: Aleutian mink disease virus

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: A36760; A35529  
R;Bloom, M.E.; Alexandersen, S.; Perryman, S.; Lechner, D.; Wolfbarger, J.B.  
J. Virol. 62, 2903-2915, 1988  
A;Title: Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus  
A;Reference number: A36760; MUID:88275062; PMID:2839709  
A;Accession: A36760  
A;Molecule type: DNA  
A;Residues: 1-620 <BL2>  
A;Cross-references: UNIPROT:P24030; EMBL:M20036  
C;Superfamily: parvovirus noncapsid protein  
C;Keywords: noncapsid protein

Query Match 14.9%; Score 543; DB 1; Length 620;  
Best Local Similarity 26.6%; Pred. No. 1.9e-31;  
Matches 157; Conservative 104; Mismatches 229; Indels 100; Gaps 18;

Qy 18 KEKSNQEVFVFVKENNVQLNGDKIGWNSYKKELOEDEKSLQGAETWTDQSE-DMEWE 76  
Db 38 EQRRLQDLVYVQLKKEIN---DGEGLAWLFQKQVTDKDNKPTKATPLRTTSSDLRLAFD 94  
Qy 77 TTVDEMTKKQVFIQSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWKGQDQGW 126  
Db 95 SIEENLTASNEHLTNNEINFCKLTGLKTLILLIDKHVSHRWDNSKNVLIWQIEKGKTOQF 154  
Qy 127 HCHVLIG-----GKDFSOAQGWRRQLN-----VYMSRWLVYACNVQ--LTPAERIKL 173  
Db 155 HIHCCILGYFDKNDPKDVQKSLG-WFMKRLNKLDAVIYSN---HHCDIQIDKDPEDRAKN 210  
Qy 174 RETAEDNEWVTLTYKHKOTKDYTKCVLFGNMIAYFLTKKISTSPRDRGCVFLSSDS 233  
Db 211 LKVMIEDGPTKPYKFNKQKDYKQVFNKPTNLRYTLIYLFNKDKINTDSM--DGYFAAGNG 269  
Qy 234 GWKTNFLKGERHLVSKLYTDDMRPETVETV-----TTAQETKRG----- 274  
Db 269 GIVDN-LTNKERTKRLKWLDEQSSDIMANDIDWEDQDAPKVTDDQSDATTKTGLSLIW 327  
Qy 275 -----RIQTKKEVA-----IKTTLKELVHKRVTSPEDEMMQSDSIEMMAQP 317  
Db 328 KSCATKVTSSKEVANVPQPSKKLYSAQSTLDALFNVGCTPEDMIKQSDKYLESLSP 387  
Qy 318 GGENLLKNTLEICTLTARTKTAPDLILE--KAETSKLTNFSLPDTRTCRIFAFHGWYV 375  
Db 388 NGPQKINTLLHMNQVKTSTMTAFDCIIFNEEBEDKPLLATIKOM-----GLNEQ 438  
Qy 376 KVCHAICCVLNRQGGKNTVLFHGPASTGKSIITAOIAQAVNGVGVNAAVNPFPNDCT 435  
Db 439 YLKKVLTILTQSGKRGCVWFYFGPGTGKTLASLICKATVNYGMVTTSNPNFPWTDG 498  
Qy 436 NKNLIIWVEBAGNFQOVNQFKAICSQGTIRIDQKGSQKQIEPTPTVIMTNNITVVIRIG 495  
Db 499 NRRNIIWABECGNFGNWNVEDFKAITGGDVKVDTKNKPQSIKGC-VIVTSNTNITKVTVG 557  
Qy 496 CEERPEHTQPIRDRMLNIHL-----THLPGDFGLVDKNEWPMICAWL 538  
Db 558 CVETNAHAEPKQRMKIRCKMTINPKTKITPG-----MLKRWL 596

## RESULT 14

S41861

gene NS-1 protein - Aleutian mink disease virus

C;Species: Aleutian mink disease virus

C;Date: 15-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C;Accession: S41861; S41436

R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.

submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of

A;Reference number: S41434

A;Accession: S41861

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-641 <GOT>

A;Cross-references: UNIPROT:Q65023; EMBL:X77085; NID:g452597; PIDN:CAA54355.1; PID:g45259

## C;Genetics:

A;Introns: 586/1

C;Superfamily: parvovirus noncapsid protein

Query Match 14.7%; Score 533.5; DB 2; Length 641;  
Best Local Similarity 25.7%; Pred. No. 1e-30;  
Matches 178; Conservative 106; Mismatches 250; Indels 159; Gaps 27;

QY 18 KEKSNQVFGVFVFNENVQLNGKDGWNSYKKEIQEDE-----LKSILQGAETWD 68  
DB 8 EQRKLQDLYQLKE---VADGEGLAWLFOQKYTTDKNRPATPLPRTSSDLRLAFD 64  
QY 69 QSEDM-----EWETTVDGMT-----KKQVFIIDSLSVKKCLFEVLNTKNIFPFGDVNWFVQ 117  
DB 65 SIBETLIASNEW-LTKDEINFCKLTGLKTLVLVDKHKVSHRWA-----DKINFIWQ 115  
QY 118 HEWGKQGWCHVLIG-----GKDFSQAQGWRRQLNYSRWMLVACNVQLTPAER 170  
DB 116 IEKGTQHFHHCGLGYFDKNDPKQVQKSLG-WLIKLN-----RDLAIFSNHHCDOOD 170  
QY 171 IKLREIAEDN--EWV-----TLTYKHQTKDKYTKCVLFGNNMIAYVFLTKKXISTSPPR 223  
DB 171 IKOPEAKANLKWIEDGPKPYKYNKQTKQYKPVLTSDVYIYLFNKXIH-----K 226  
QY 224 DG--GYPLSDSGWKNFLKEGERHLVSKLYTDDMRPETVE-----TTVTT 267  
DB 227 EGMGGYAAAGGLIDN-LTNKERRKALRWYLDQSSDIMDADIDWEDGQDAPKVDQTD 285  
QY 268 AQETKRG-----RIOTKKEVA-----IKTTLKELVHKVTSPEWMM 305  
DB 286 SATSKTGTSLIWKSCATKVTSKVEANPVQOPSKKLYSAQNTLDALFNVCFTPEDMIK 345  
QY 306 QPDSYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILE--KAETSKLTNFSLPDTRT 363  
DB 346 QSDKYLELSELPNGPKINTLLHNVQKYSTMTATFDCLIKFNEEDDKPLATIIDM-- 403  
QY 364 CRIFAFHGMNYSVCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAVGNVCYN 423  
DB 404 -----GLNEQYLKVKVLTILTQGGKRCIWFYFGGTGKTLASLICKATVNYGWT 456  
QY 424 AANVPFNDCTNKLWEEAGNFQGVNQFKAICSGQTIRIDQKGSQKIEPTPVIM 483  
DB 457 TSNPNFPWTDCGRNIWAECCNLGNWVEDFKAITGGGDKVYDTKKNQPSIKGC-VTV 515  
QY 484 TTINENITVVRIGCEERPEHTQPIRDRMLNIHL-----THLPDGFGLVDKNEWPMICA 536  
DB 516 TSNNTITKVTGCVETNAHAELPKQRMKIRCMKTNPKTKITPG-----MLKR 564  
QY 537 WLKNGVQSTMASYCAKWKVP-----DWSNWAEPKVPTPINLLGSARSPF-TTPKSTP 590  
DB 565 WL-----NTWDRQPTQLSHEMPELYLETTGPN-----SSATSATKTTGNSQP 606  
QY 591 LSQNYALTPLASDLLELPWSTPNTPVAGTA 623  
DB 607 TTAKTAEVSSTADCD-----TPKRKAS 628

## RESULT 15

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N;Alternate names: nonstructural protein NS1; NS1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S52209

R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52209

A;Molecule type: DNA

A;Residues: 1-626 &lt;ZAD&gt;

A;Cross-references: UNIPROT:O83288; EMBL:X75093; NID:g609091; PIDN:CAA52983.1; PID:g60909

A;Experimental source: strain FM

C;Genetics:

## A;Gene: NS1

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 12.4%; Score 452.5; DB 2; Length 626;  
Best Local Similarity 25.2%; Pred. No. 7.5e-25;  
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKQVFIIDSLSVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
DB 45 WBEPTGIWNNEHVNLPMTLADKI-----KNIFIQRWNQFNQDSTDFFFQLEEGS 93  
QY 123 DQGWCHVLIGKDFSQAQGWRRQLNYSRWMLVACNVQLTPAERIK--LREIAD 179  
DB 94 EY-IHLHACVPGECRSFVLGRY-----MSQIKDILRDVYEG 129  
QY 180 N-----EWTLTYKHQTKDKYTKCVLFGNNMIAYVFLTKKXISTSPPRDGGYFSSDSG 234  
DB 130 KQVKIPDMFSITKTKRGGQNKTVTAA-----YILHYLIPKKQ-----PELOWAFNTMPLF 179  
QY 235 WKTNFKGERHLVSKLYTDDMRPETVETVTTAQETKEGRIOTKKEVAIKTTLKELVHK 294  
DB 180 TTAALCLQKQRELLDAFQSENNNAVQEDQAATAPL-----ISNRAAKNYSNLVDWLIEM 235  
QY 295 RVTSPEDMMWOPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354  
DB 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295  
QY 355 NPSLPDTRTCRIFAFHGMNYSVCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQVGVLCGWVKRBEFNKRNAILWYGPATTTGKTNIAEATAH 350  
QY 415 AVGNVCYNAANVPENPDCTNKLWEEAGNFQGVNQFKAICSGQTIRIDQKGSK 474  
DB 351 AVPFYGCNVNWTNENPFNDVDKMLIWEEGKNTKRVESAKAILGGSARVDDQCKGSV 410  
QY 475 QIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGFGLVDKNE---- 530  
DB 411 CIEPTPVITSTNDMCMIVDGNSTTMEHRIPLERMFQIVLSHKLGNFGKISKEVKEF 470  
QY 531 --W-----PMTCAWLKNGYQSTWASCAKWKGVDPDWSN-----WAEPK----- 568  
DB 471 FKWANDNLVPVSEPKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPTRBELBEI 524  
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLLELPWSTPNT 617  
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGETQVRCAMHSL-DNSMNVFECLEC 576  
QY 618 PVAGTAETQNTGEAGSKAC-QDQGLSPTWSRIEEDLRACFGAEPPLKDPFSE 667  
DB 577 ERANPPEFQSLGE---NFCNQHG-----WYD-----CAFCELNKDDMNE 612

Search completed: January 22, 2005, 03:41:57

Job time : 26.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:19:28 ; Search time 106.333 Seconds  
(without alignments)  
3636.223 Million cell updates/sec

Title: US-10-069-056-6  
Perfect score: 3638  
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3625	99.6	672	1	VNCS_MUMIV	P03134 murine minu
2	3625	99.6	721	2	Q84365	Q84365 murine minu
3	3556	97.7	721	1	VNCS_MUMIM	P07300 murine minu
4	3556	97.7	721	2	Q84363	Q84363 murine minu
5	3554	97.7	672	2	Q83429	Q83429 mouse parvo
6	3347	92.0	672	2	Q8JV14	Q8JV14 rat minute
7	3338	91.8	672	2	Q8JV28	Q8JV28 kilham rat
8	3337	91.7	672	2	P88999	P88999 kilham rat
9	3329	91.5	672	2	Q8JV18	Q8JV18 rat minute
10	3328	91.5	672	1	VNCS_PAVHH	P03133 hamster par
11	3325	91.4	672	2	Q8JV16	Q8JV16 rat minute
12	3324	91.4	668	1	VNCS_PAVL3	P36311 parvovirus
13	3305	90.8	665	2	O71159	O71159 kilham rat
14	2937.5	80.7	671	2	O71157	O71157 rat parvovi
15	2657	73.0	668	2	P89513	P89513 feline panl
16	2656	73.0	668	2	P89516	P89516 feline panl
17	2655	73.0	668	2	P89512	P89512 feline panl
18	2655	73.0	668	2	P89515	P89515 feline panl
19	2655	73.0	668	2	P90449	P90449 feline panl
20	2655	73.0	668	2	P90484	P90484 feline panl
21	2654	73.0	668	2	P89514	P89514 feline panl
22	2654	73.0	668	2	P90472	P90472 feline panl
23	2649	72.8	668	1	VNCS_MEVA	P27438 mink enteri
24	2646	72.7	668	1	VNCS_FPV19	P24842 feline panl
25	2644	72.7	668	1	VNCS_PAVCN	P12929 canine parv
26	2639	72.5	668	2	Q84393	Q84393 canine parv
27	2616	71.9	668	2	Q70M74	Q70M74 canine parv
28	2615	71.9	668	2	CAE47433	CAE47433 canine pa
29	2487	68.4	660	1	VNCS_PAVPN	P18547 porcine par
30	2478	68.1	662	1	VNCS_PAVPK	P52502 porcine par
31	2474	68.0	662	2	Q6RED5	Q6RED5 porcine par

RESULT 1

ID	VNCS_MUMIV	STANDARD;	PRT;	672 AA.
AC	P03134;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).			
GN	Name=NS1;			
OS	Murine minute virus (Murine parvovirus).			
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.			
OX	NCBI_TaxID=10794;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=83143341; PubMed=6298737;			
RA	Astell C.R., Thomson M., Merchinsky M., Ward D.C.;			
RT	"The complete DNA sequence of minute virus of mice, an autonomous parvovirus.";			
RL	Nucleic Acids Res. 11:999-1018(1983).			
RN	[2]			
RP	HOMOLOGIZATION, AND INTERACTION WITH SYNCRIP.			
RP	MEDLINE=99102562; PubMed=9847309;			
RA	Harris C.E., Boden R.A., Astell C.R.;			
RT	"A novel heterogeneous nuclear ribonucleoprotein-like protein interacts with NS1 of the minute virus of mice.";			
RL	J. Virol. 73:72-80(1999).			
CC	-I- FUNCTION: Seems necessary for viral DNA replication.			
CC	-I- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP.			
CC	-I- DOMAIN: The N-terminus (residues 1-275) possess a negative effect on transactivation.			
CC	-I- DOMAIN: The C-terminus (residues 543-672) possess an activation domain.			
CC	-I- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; J02275; AAA67109.1; -			
DR	EMBL; V01115; CAA24309.1; ALT_INIT.			
DR	PIR; A03696; UYPVIM.			
DR	TRANSFAC; T02375; -			
DR	InterPro; IPR001257; Parvo NS1.			
DR	Pfam; PF01057; Parvo NS1; I.			
KW	ATP-binding; DNA replication; Noncapsid protein;			
KW	Nonstructural protein.			
FT	DOMAIN 1 276 Interacts with SYNCRIP.			
FT	NP_BIND 399 406 ATP (Potential)			
SQ	SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;			

O6TPD8 porcine par  
Aa90279 porcine p  
Aa91039 porcine p  
Q6PS60 porcine par  
Aa93262 porcine p  
Q993m6 autonomous  
P06431 feline panl  
Q65017 aleutian mi  
Q65020 aleutian mi  
P24030 aleutian mi  
Q96607 aleutian mi  
Q65023 aleutian mi  
Q6j180 bovine aden  
Aa26464 bovine ad

```
Query Match 99.6%; Score 3625; DB 1; Length 672;
Best Local Similarity 99.7%; Pred. No. 1.e-260;
Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVIGATNWLKESNQEVFSFVKNNVQNGKDIGWNSYKKELOEDELKSLQ 60
DB 1 MAGNAYSDEVIGATNWLKESNQEVFSFVKNNVQNGKDIGWNSYKKELOEDELKSLQ 60
QY 61 RGAETTTWQSDMEWETTVDEMTKKQVIFDLSVKKCLFEVLNTKNIFFPGDVNVFVQHEW 120
DB 61 RGAETTTWQSDMEWETTVDEMTKKQVIFDLSVKKCLFEVLNTKNIFFPGDVNVFVQHEW 120
QY 121 GKDGQWHCHVLIGKDFSAQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB 121 GKDGQWHCHVLIGKDFSAQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHQTKDQYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
DB 181 EWTLLTYKHQTKDQYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKRGRIQTKKEVAIKTKLKHVKRVTSP 300
DB 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKRGRIQTKKEVAIKTKLKHVKRVTSP 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
DB 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
QY 361 TRTCRIFAFHGNVYKVAICCVLNROGGKNTVLFHGPASTGKSIIAQAIAQAVGNVG 420
DB 361 TRTCRIFAFHGNVYKVAICCVLNROGGKNTVLFHGPASTGKSIIAQAIAQAVGNVG 420
QY 421 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
DB 421 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
QY 541 NGYSTWASVYCAKWKVPDSENWAEKVPPTINLLGSARSPTTTPKSTPLSONYALTPL 600
DB 541 NGYSTWASVYCAKWKVPDSENWAEKVPPTINLLGSARSPTTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
DB 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 2
Q84365 PRELIMINARY; PRT; 721 AA.
AC Q84365;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NSI;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MVM;
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus."
RL Nucleic Acids Res. 11:999-1018(1983).
```

```
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=MVM;
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MYM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain."
RL J. Virol. 57:656-669(1986).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=MVM;
RX MEDLINE=87061199; PubMed=3783817;
RA Morgan W.R., Ward D.C.;
RT "Three splicing patterns are used to excise the small intron common to
RT all minute virus of mice RNAs."
RL J. Virol. 60:1170-1174(1986).
DR EMBL; J02275; AAA67108.1; -.
DR GO; GO:0019012; C:viral genome replication; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo.NS1.
DR Pfam; PF01057; Parvo.NS1.
KW Nonstructural protein.
SQ SEQUENCE 721 AA; 81896 MW; 18391758E42F0DCF CRC64;

Query Match 99.6%; Score 3625; DB 2; Length 721;
Best Local Similarity 99.7%; Pred. No. 1.e-260;
Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVIGATNWLKESNQEVFSFVKNNVQNGKDIGWNSYKKELOEDELKSLQ 60
DB 50 MAGNAYSDEVIGATNWLKESNQEVFSFVKNNVQNGKDIGWNSYKKELOEDELKSLQ 109
QY 61 RGAETTTWQSDMEWETTVDEMTKKQVIFDLSVKKCLFEVLNTKNIFFPGDVNVFVQHEW 120
DB 110 RGAETTTWQSDMEWETTVDEMTKKQVIFDLSVKKCLFEVLNTKNIFFPGDVNVFVQHEW 169
QY 121 GKDGQWHCHVLIGKDFSAQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB 170 GKDGQWHCHVLIGKDFSAQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLREIAEDN 229
QY 181 EWTLLTYKHQTKDQYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
DB 230 EWTLLTYKHQTKDQYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 289
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKRGRIQTKKEVAIKTKLKHVKRVTSP 300
DB 290 KEGERHLVSKLYTDDMRPETVETTTAQTETKRGRIQTKKEVAIKTKLKHVKRVTSP 349
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
DB 350 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 409
QY 361 TRTCRIFAFHGNVYKVAICCVLNROGGKNTVLFHGPASTGKSIIAQAIAQAVGNVG 420
DB 410 TRTCRIFAFHGNVYKVAICCVLNROGGKNTVLFHGPASTGKSIIAQAIAQAVGNVG 469
QY 421 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
DB 470 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGSQKQIEPTP 529
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
DB 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 589
QY 541 NGYSTWASVYCAKWKVPDSENWAEKVPPTINLLGSARSPTTTPKSTPLSONYALTPL 600
DB 590 NGYSTWASVYCAKWKVPDSENWAEKVPPTINLLGSARSPTTTPKSTPLSONYALTPL 649
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
DB 650 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 709
```



```

Qy 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 3
VNC5 MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Name=NS1;
OS Murine minute virus (strain MVM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP (By
CC similarity).
CC -!- DOMAIN: The N-terminus (residues 1-275) possess a negative effect
CC on transactivation (By similarity).
CC -!- DOMAIN: The C-terminus (residues 543-672) possess an activation
CC domain (By similarity).
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02481; -; NOT_ANNOTATED_CDS.
DR EMBL; M12032; AAA69567.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT DOMAIN 1 276 Interacts with SYNCRIP (By similarity).
FT NP_BIND 399 406 ATP (Potential).
FT FT CONFLICT 597 597 I -> L (in Ref. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 97.7%; Score 3556; DB 1; Length 672;
Best Local Similarity 97.5%; Pred. No. 1.4e-255;
Matches 655; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVILGATNMLKESNOEVSFVPEKPNVOLNGDKIGWNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVILGATNMLKESNOEVSFVPEKPNVOLNGDKIGWNSYKKELODELKSLQ 60
Qy 61 RGAETTWQSQEDMEWETTVDMTKKQVFIPIFDSLVKKCLFEVLNKNIFPGDVNWFVQHEW 120
Db 61 RGAETTWQSQEDMEWETTVDMTKKQVFIPIFDSLVKKCLFEVLNKNIFPGDVNWFVQHEW 120
Qy 121 GKDQGHCHVLIGGKDFSOAQGKWRRLQNLVYWSRWLVLTACNVQLTPAERIKUREIAEDN 180

```

```
Qy 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVLQNGKDIGNWSYKKELOEDELKSLQ 60
Db 50 MAGNAYSDEVLTGTTNWLKEKSNQEVFSFVKETDVLQNGKDIGNWNKYKKELOEDELKSLQ 109
Qy 61 RGAETTWDSQSEDMWETTVDEMTHKQVFIPIFDSLKKCLFEVLNTKNIIPFGDVNWFVQHEW 120
Db 110 RGAETTWDSQSEDMWETTVDEMTHKQVFIPIFDSLKKCLFEVLNTKNIIPADVTWVQHEW 169
Qy 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVMYWSRWLVTAQNVOLTPAERIKLREIAEDN 180
Db 170 GKDGQWHCHVLIGGKDFSAQGWRRQLNVMYWSRWLVTAQNVOLTPAERIKLREIAEDS 229
Qy 181 EWTLLTYKHQTKQDYTKCVLFGNMIAIYFLTKKKISTSPRRDGGYFLSSDGSWKTNFL 240
Db 230 EWTLLTYKHQTKQDYTKCVLFGNMIAIYFLTKKKISTSPRRDGGYFLSSDGSWKTNFL 289
Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
Db 290 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 349
Qy 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAEISKLTNFSPLD 360
Db 350 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAEISKLTNFSPLD 409
Qy 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIQAAGVNGV 420
Db 410 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIQAAGVNGV 469
Qy 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGKSKQIEPTP 480
Db 470 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGKSKQIEPTP 529
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 589
Qy 541 NGYSTWASACAKGKVPDWSNENWAEKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
Db 590 NGYSTWASACAKGKVPDWSNENWAEKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 649
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 660
Db 650 ASDLEDLALPWSPTNPVAGTAETQNTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 709
Qy 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 5
Q83429 ID Q83429 PRELIMINARY; PRT; 672 AA.
AC Q83429;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Names=NS1;
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP MEDLINE=94365951; PubMed=8081985;
RX Ball-Goodrich L.J., Johnson E.;
RT "Molecular characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486(1994).
DR EMBL; U12469; AAA61405.1; --
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
KW Nonstructural protein.
```

```
SQ SEQUENCE 672 AA; 76112 MW; 31CG365276727363 CRC64;
Query Match 97.7%; Score 3554; DB 2; Length 672;
Best Local Similarity 97.3%; Pred. No. 1.9e-255;
Matches 654; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVLQNGKDIGNWSYKKELOEDELKSLQ 60
Db 1 MAGNAYSDEVLTGTTNWLKEKSNQEVFSFVKETDVLQNGKDIGNWNKYKKELOEDELKSLQ 60
Qy 61 RGAETTWDSQSEDMWETTVDEMTHKQVFIPIFDSLKKCLFEVLNTKNIIPFGDVNWFVQHEW 120
Db 61 RGAETTWDSQSEDMWETTVDEMTHKQVFIPIFDSLKKCLFEVLNTKNIIPADVTWVQHEW 120
Qy 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVMYWSRWLVTAQNVOLTPAERIKLREIAEDN 180
Db 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVMYWSRWLVTAQNVOLTPAERIKLREIAEDS 180
Qy 181 EWTLLTYKHQTKQDYTKCVLFGNMIAIYFLTKKKISTSPRRDGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKQDYTKCVLFGNMIAIYFLTKKKISTSPRRDGGYFLSSDGSWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
Qy 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAEISKLTNFSPLD 360
Db 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAEISKLTNFSPLD 360
Qy 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIQAAGVNGV 420
Db 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIQAAGVNGV 420
Qy 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGKSKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGKSKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
Qy 541 NGYSTWASACAKGKVPDWSNENWAEKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
Db 541 NGYSTWASACAKGKVPDWSNENWAEKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAETQNTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 660
Qy 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 6
Q8JV14 ID Q8JV14 PRELIMINARY; PRT; 672 AA.
AC Q8JV14;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NS1;
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat parvoviruses.";
KW parvoviruses."
```

```
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF332884; AAM93279.1; -.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:0019079; P: viral genome replication; IEA.
DR InterPro: IPR001257; Parvo_NSL.
DR Pfam: PF01057; Parvo_NSL; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 92.0%; Score 3347; DB 2; Length 672;
Best Local Similarity 91.7%; Pred. No. 4.9e-240;
Matches 616; Conservative 23; Mismatches 33; Indels 0; Gaps 0;

QY 1 MAGNAYSDVLGATNWLKESNQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
DB 1 MAGNAYSDVLGATNWLKESQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
QY 61 RGAETTWQSDSEMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
DB 61 RGAETTWQSDSEMEWESA VDDMTKQVFI FDSL VKKCLFEVLSTKNI APSDVTWVQHEW 120
QY 121 GKQGHCHVLIGKDFSOAQGWRRQLNVTYNSRWLVTA CNVQLTPAERI KILREIAEDN 180
DB 121 GKQGHCHVLIGKDFSOAQGWRRQLNVTYNSRWLVTA CNVQLTPAERI KILREIAEDS 180
QY 181 EWTLLTYKHQTKDYTKCVLFGNMIAYVFLTKKISTSPR DGGVFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHHTKDYTKCVLFGNMIAYVFLSKKICTSPR DGGVFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMPETVETTVTAQEA KGRIOTRKEVSIKTTLLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAS 360
QY 361 TRTCRIFAEHGMNVIKCHAI CCVLRNQGKRNVTLPFGHPASTGKSI IAAIAQAGVNVG 420
DB 361 TRTCRIFAEHGMNVIKCHAI CCVLRNQGKRNVTLPFGHPASTGKSI IAAIAQAGVNVG 420
QY 421 CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
DB 421 CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDSENAEPKVPTPINLIGSARSPPTPKSTPLSONYALTPL 600
DB 541 NGYQSTMASYCAKWKVPDSENAEPKVPTPINLIGSARSPPTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEPWTPTNTPVAGTAETQNTGEAGSKACODGQLSPTWSEIEDLRACFCAEP 660
DB 601 ASDLEDLALPEPWTPTNTPVAGTAETQNTGEAGSKACODGQLSPTWSEIEDLRACFCAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 7
Q8JV28 PRELIMINARY; PRT; 672 AA.
ID AC
IC Q8JV28:
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Nonstructural protein 1.
GN Name=ns1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses.";
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Soderlund-Venermo M., Pintel D., Riley L.K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF321230; AAM93272.1; -.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:0019079; P: viral genome replication; IEA.
DR InterPro: IPR001257; Parvo_NSL.
DR Pfam: PF01057; Parvo_NSL; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75987 MW; 22B4611C20CDB6E9 CRC64;

Query Match 91.8%; Score 3338; DB 2; Length 672;
Best Local Similarity 91.5%; Pred. No. 2.3e-239;
Matches 615; Conservative 23; Mismatches 34; Indels 0; Gaps 0;

QY 1 MAGNAYSDVLGATNWLKESNQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
DB 1 MAGNAYSDVLGATNWLKESQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
QY 61 RGAETTWQSDSEMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
DB 61 RGAETTWQSDSEMEWESA VDDMTKQVFI FDSL VKKCLFEVLSTKNI APSDVTWVQHEW 120
QY 121 GKQGHCHVLIGKDFSOAQGWRRQLNVTYNSRWLVTA CNVQLTPAERI KILREIAEDN 180
DB 121 GKQGHCHVLIGKDFSOAQGWRRQLNVTYNSRWLVTA CNVQLTPAERI KILREIAEDS 180
QY 181 EWTLLTYKHQTKDYTKCVLFGNMIAYVFLTKKISTSPR DGGVFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHHTKDYTKCVLFGNMIAYVFLSKKICTSPR DGGVFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMPETVETTVTAQEA KGRIOTRKEVSIKTTLLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAS 360
QY 361 TRTCRIFAEHGMNVIKCHAI CCVLRNQGKRNVTLPFGHPASTGKSI IAAIAQAGVNVG 420
DB 361 TRTCRIFAEHGMNVIKCHAI CCVLRNQGKRNVTLPFGHPASTGKSI IAAIAQAGVNVG 420
QY 421 CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
DB 421 CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDSENAEPKVPTPINLIGSARSPPTPKSTPLSONYALTPL 600
DB 541 NGYQSTMASYCAKWKVPDSENAEPKVPTPINLIGSARSPPTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEPWTPTNTPVAGTAETQNTGEAGSKACODGQLSPTWSEIEDLRACFCAEP 660
DB 601 ASDLEDLALPEPWTPTNTPVAGTAETQNTGEAGSKACODGQLSPTWSEIEDLRACFCAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672
```

```
Db      661 LKDFSEPLNLD 672
       1:||||:|
       661 LESDFNEELTLD 672
       1:||||:|

RESULT 8
P88899 PRELIMINARY; PRT; 672 AA.
AC P88899;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Non-capsid protein.
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown D.W., Like A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U79033; AAB38326.1; -.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match 91.7%; Score 3337; DB 2; Length 672;
Best Local Similarity 91.4%; Pred. No. 2.7e-239;
Matches 614; Conservative 23; Mismatches 35; Indels 0; Gaps 0;

Qy 1 MAGNAYSDVLGATNWLKESNQVFSFVKFNENVQLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDVLGATNWLKSSQEVFSFVKFNENVQLNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTTWQSDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Db 61 RGAETTTWQSDMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Qy 121 GKDGQWHCHVLIGKDFSQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLRETAEDN 180
Db 121 GKDGQWHCHVLIGKDFSQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLRETAEDS 180
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDRGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLSKKICTSPRDRGGYFLSSDGSWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPEVETTTAQTKEGRIOTKKEVAIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPEVETTTAQTKEGRIOTKKEVAIKTTLKELVHKRVTSPE 300
Qy 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLILEKAETSKLTNFSLPD 360
Db 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLILEKAETSKLANFSMAN 360
Qy 361 TRTCRIFAFHGNYKVKCHAI CCVLNRQGGKRNVL FPHGPASTGKSI IIAQIAQAVGNVG 420
Db 361 TRTCRIFAFHGNYKVKCHAI CCVLNRQGGKRNVL FPHGPASTGKSI IIAQIAQAVGNVG 420
Qy 421 CYNAAVNVPNDCTNKNLIWVEEAGNFQGVNQFKAICSGOTIRIDQKKGSKQIEPTP 480
Db 421 CYNAAVNVPNDCTNKNLIWVEEAGNFQGVNQFKAICSGOTIRIDQKKGSKQIEPTP 480

Query Match 91.5%; Score 3329; DB 2; Length 672;
Best Local Similarity 91.2%; Pred. No. 1.1e-238;
Matches 613; Conservative 24; Mismatches 35; Indels 0; Gaps 0;

Qy 1 MAGNAYSDVLGATNWLKESNQVFSFVKFNENVQLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDVLGATNWLKSSQEVFSFVKFNENVQLNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTTWQSDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Db 61 RGAETTTWQSDMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Qy 121 GKDGQWHCHVLIGKDFSQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLRETAEDN 180
Db 121 GKDGQWHCHVLIGKDFSQAGKWRRLQNVYWSRWLVTAACNVQLTPAERIKLRETAEDS 180
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDRGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLSKKICTSPRDRGGYFLSSDGSWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPEVETTTAQTKEGRIOTKKEVAIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPEVETTTAQTKEGRIOTKKEVAIKTTLKELVHKRVTSPE 300
Qy 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLILEKAETSKLTNFSLPD 360
Db 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLILEKAETSKLANFSMAN 360
Qy 361 TRTCRIFAFHGNYKVKCHAI CCVLNRQGGKRNVL FPHGPASTGKSI IIAQIAQAVGNVG 420
Db 361 TRTCRIFAFHGNYKVKCHAI CCVLNRQGGKRNVL FPHGPASTGKSI IIAQIAQAVGNVG 420
Qy 421 CYNAAVNVPNDCTNKNLIWVEEAGNFQGVNQFKAICSGOTIRIDQKKGSKQIEPTP 480
Db 421 CYNAAVNVPNDCTNKNLIWVEEAGNFQGVNQFKAICSGOTIRIDQKKGSKQIEPTP 480
```

```

QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKHEWPLICAWLVK 540
QY 541 NGYQSTMASYCAKMGKVPDWSNWAEPKVPPTINLLGSARSPTTPKSTPLSONYALTPL 600
DB 541 NGYQSTMACYCAKMGKVPDWSNWAEPKLETPINSIGSMRSPSLTPRSTPLSONYALTPL 600
QY 601 ASDLELALPEWSTPNTVPVAGTAEONTQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
DB 601 ASDLADLALPEWSTPNTVPVAGTAAQSNTGAGSTACQGAQRSPTWSEIEADLRACFSQEQ 660
QY 661 LKDFSEPLNLD 672
DB 661 LKDFSDSLTLD 672

RESULT 10
VNC5_PAVH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Names=NS1;
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184 (1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X01457; CA925689.1; --
DR PIR; A03695; UTPV1.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT NP_BIND 399 406 ATP (Potential).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 91.5%; Score 3328; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 1.3e-238;
Matches 612; Conservative 25; Mismatches 35; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKESKNSQEVFSFVFPKENVQLNGKDIWNYSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVILGVTNMLKDKSSQEVFSFVFPKENVQLNGKDIWNYSYKKELODELKSLQ 60
QY 61 RGAETTWDSQEDMEWETTVDEMTPKQVIFPDSLVKKCLFVLNTKNIIFGCVNMFVQHEW 120
DB 61 RGAETTWDSQEDMEWESAVDDMTKQVIFPDSLVKKCLFVLNTKNIAPSNVTFVQHEW 120
QY 121 GKQDGHCHVLIGKDFPSQAGKWRRLNVLNYSRMLVTACNVQLTPAERIKLREIAEDN 180
DB 121 GKDPGHCHVLIGKDFPSQAGKWRRLNVLNYSRMLVTACNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHKTKDYTKCVLFGNMIAIYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240

```

```

DB 181 EWTLLTYKHKTKDYTKCVLFGNMIAIYFLSKKICTSPPRDGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMKEPETVETVTTAQEAQKRGRIQTRREEVSIKTTKLKELVHKRVTSPE 300
QY 301 DWMQOPDSYIEEMQAOPGGENLLKNTLEICTLTARTKTAFDLILKKAETSKLTNFSLPD 360
DB 301 DWMQOPDSYIEEMQAOPGGENLLKNTLEICTLTARTKTAFDLILKKAETSKLANFSMAS 360
QY 361 TTTCTRIFAHHGNNYKVCCHAI CCVLNRQGGKNTVLFHGPASTGKSI IAQAIAQAAGNVG 420
DB 361 TTTCTRIFAHHGNNYIKVCHAI CCVLNRQGGKNTVLFHGPASTGKSI IAQAIAQAAGNVG 420
QY 421 CYNAAANVFPFNDCTNKNLIWVEEAGNFQGVQVQNFKAICSGQTIRIDQKKGSKQIEPTP 480
DB 421 CYNAAANVFPFNDCTNKNLIWVEEAGNFQGVQVQNFKAICSGQTIRIDQKKGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKHEWPLICAWLVK 540
QY 541 NGYQSTMASYCAKMGKVPDWSNWAEPKVPPTINLLGSARSPTTPKSTPLSONYALTPL 600
DB 541 NGYQSTMACYCAKMGKVPDWSNWAEPKLDTPINSIGSMRSPSLTPRSTPLSONYALTPL 600
QY 601 ASDLELALPEWSTPNTVPVAGTAEONTQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
DB 601 ASDLADLALPEWSTPNTVPVAGTAAQSNTGAGSTACQGAQRSPTWSEIEADLRACFSQEQ 660
QY 661 LKDFSEPLNLD 672
DB 661 LKDFSEPLTLD 672

RESULT 11
QY 661 PRELIMINARY; PRT; 672 AA.
ID Q8JVL6
AC Q8JVL6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Names=NS1;
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses.";
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332883; AM93277.1; --
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71F6BEF449A6 CRC64;

Query Match 91.4%; Score 3325; DB 2; Length 672;
Best Local Similarity 90.9%; Pred. No. 2.1e-238;
Matches 611; Conservative 25; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKESKNSQEVFSFVFPKENVQLNGKDIWNYSYKKELODELKSLQ 60

```

Db 1 MAGNAYSDEVLGATNWLKDKSQSEVFSFVFNENVLQNGKDIGNSYRKLQDELKSLQ 60  
Qy 61 RGAETTWDSQSEDMETTTVDKTKQVFIPLSLVKKCLFEVLNTKNIIPGQVNVFWQHEW 120  
Db 61 RGAETTWDSQSEDMESAVDDMTKKQVFIPLSLVKKCLFEVLNTKNIIPGQVNVFWQHEW 120  
Qy 121 GKDGQWHCHVLIGGKDFSAQGWRRQNLVYWSRWLVATCNVQLTPAERIKLREIAEDN 180  
Db 121 GKDGQWHCHVLIGGKDFSAQGWRRQNLVYWSRWLVATCNVQLTPAERIKLREIAEDS 180  
Qy 181 EWTLLTYKHQTKQVTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKNTNPL 240  
Db 181 EWTLLTYKHQTKQVTKCVLFGNMIAYFLSKKIKTSPRGGYFLSSDSGKNTNPL 240  
Qy 241 KEGERHLVSKLYTDDMPETVTTTAQETKRGRIOTKKEVALKTLKELVHKRVTSPE 300  
Db 241 KEGERHLVSKLYTDDMPETVTTTAQEAAGRRIOTRKEVSIKTLKELVHKRVTSPE 300  
Qy 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPLSD 360  
Db 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPLSD 360  
Qy 361 TRTCRIFAFHGWNYKVCCHAICCVLNROGKRNVTLPFGPASTGKSIIAQAIAQAVNGV 420  
Db 361 TRTCRIFAFHGWNYKVCCHAICCVLNROGKRNVTLPFGPASTGKSIIAQAIAQAVNGV 420  
Qy 421 CYNAAVNFPPNDCTNKNLIWBEAGNFQGVNQFKALCSGQTRIDQKGSQKQIEPTP 480  
Db 421 CYNAAVNFPPNDCTNKNLIWBEAGNFQGVNQFKALCSGQTRIDQKGSQKQIEPTP 480  
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540  
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540  
Qy 541 NGYQSTWASYCAKWKGVDPDSENAEKPVTPIINLGSARSPPTPKSTPLSQNYALTPL 600  
Db 541 NGYQSTWASYCAKWKGVDPDSENAEKPVTPIINLGSARSPPTPKSTPLSQNYALTPL 600  
Qy 601 ASDLEDLALPWSPTNPVAGTAEONTGAGSKACODGQSLSPTWSEIEDLRACFGAEP 660  
Db 601 ASDLEDLALPWSPTNPVAGTAEONTGAGTACQAGQARSPTWSEIEDLRACFSQEQ 660

## RESULT 12

VNCS\_PAVL3 STANDARD; PRT; 668 AA.  
AC P36311;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Noncapsid protein NS-1 (Nonstructural protein NS1).  
GN Name=NS1;  
OS Parvovirus LuIII.  
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
OX NCBI\_TaxID=35339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93297126; PubMed=8517025;  
RA Diffot N., Chen K.C., Bates R.C., Lederma M.;  
RT "The complete nucleotide sequence of parvovirus LuIII and localization  
of a unique sequence possibly responsible for its encapsidation  
pattern.";  
RT Virology 192:339-345(1993).  
RC -!- FUNCTION: Seems necessary for viral DNA replication.  
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M81888; -; NOT ANNOTATED\_CDS.  
DR PIR; A44276; A44276.  
DR InterPro; IPR001257; Parvo\_Ns1.  
DR Pfam; PF01057; Parvo\_Ns1; 1.  
KW ATP-binding; DNA replication; Noncapsid protein;  
KW Nonstructural protein.  
FT NP BIND 399 406 ATP (Potential).  
SQ SEQUENCE 668 AA; 75846 MW; CAB69049F8F86B53 CRC64;  
  
Query Match 91.4%; Score 3324; DB 1; Length 668;  
Best Local Similarity 91.4%; Pred. No. 2.5e-238;  
Matches 614; Conservative 24; Mismatches 30; Indels 4; Gaps 2;  
  
Qy 1 MAGNAYSDEVLGATNWLKDKSQSEVFSFVFNENVLQNGKDIGNSYRKLQDELKSLQ 60  
Db 1 MAGNAYSDEVLGATNWLKDKSQSEVFSFVFNENVLQNGKDIGNSYRKLQDELKSLQ 60  
Qy 61 RGAETTWDSQSEDMETTTVDKTKQVFIPLSLVKKCLFEVLNTKNIIPGQVNVFWQHEW 120  
Db 61 RGAETTWDSQSEDMESAVDDMTKKQVFIPLSLVKKCLFEVLNTKNIIPGQVNVFWQHEW 120  
Qy 121 GKDGQWHCHVLIGGKDFSAQGWRRQNLVYWSRWLVATCNVQLTPAERIKLREIAEDN 180  
Db 121 GKDGQWHCHVLIGGKDFSAQGWRRQNLVYWSRWLVATCNVQLTPAERIKLREIAEDQ 180  
Qy 181 EWTLLTYKHQTKQVTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKNTNPL 240  
Db 181 EWTLLTYKHQTKQVTKCVLFGNMIAYFLSKKIKTSPRGGYFLSSDSGKNTNPL 240  
Qy 241 KEGERHLVSKLYTDDMPETVTTTAQETKRGRIOTKKEVALKTLKELVHKRVTSPE 300  
Db 241 KEGERHLVSKLYTDDMPETVTTTAQEAAGRRIOTRKEVSIKTLKELVHKRVTSPE 300  
Qy 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPLSD 360  
Db 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPLSD 360  
Qy 361 TRTCRIFAFHGWNYKVCCHAICCVLNROGKRNVTLPFGPASTGKSIIAQAIAQAVNGV 420  
Db 361 TRTCRIFAFHGWNYKVCCHAICCVLNROGKRNVTLPFGPASTGKSIIAQAIAQAVNGV 420  
Qy 421 CYNAAVNFPPNDCTNKNLIWBEAGNFQGVNQFKALCSGQTRIDQKGSQKQIEPTP 480  
Db 421 CYNAAVNFPPNDCTNKNLIWBEAGNFQGVNQFKALCSGQTRIDQKGSQKQIEPTP 480  
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540  
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540  
Qy 541 NGYQSTWASYCAKWKGVDPDSENAEKPVTPIINLGSARSPPTPKSTPLSQNYALTPL 600  
Db 541 NGYQSTWASYCAKWKGVDPDSENAEKPVTPIINLGSARSPPTPKSTPLSQNYALTPL 600  
Qy 601 ASDLEDLALPWSPTNPVAGTAEONTGAGSKACODGQSLSPTWSEIEDLRACFGAEP 660  
Db 598 -SDLEDLALPWSPTNPVAGTAEONTGAGTACQAGQARSPTWSEIEDLRACFSSEH 656  
Qy 661 LKPDFSEPLNLD 672  
Db 657 WKSDESEQLPNLD 668

## RESULT 13

O71159  
ID O71159 PRELIMINARY; PRT; 665 AA.  
AC O71159;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds  
(without alignments)  
11177.792 Million cell updates/sec

Title: US-10-069-056-8  
Perfect score: 2019  
Sequence: 1 atggctggaatgcttactc.....agccgctgaacttggaactaa 2019

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2019	100.0	2019	6	AX137743	Sequence
2	2017.4	99.9	2019	6	AX137736	Sequence
3	2017.4	99.9	5081	14	PAMVM2	V01115 Minute viru
4	2017.4	99.9	5149	14	MVMPCG	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137739	Sequence
6	2015.8	99.8	2019	6	AX137747	Sequence
7	2015.8	99.8	2019	6	AX137751	Sequence
8	1905.4	94.4	5085	14	MVMICG	M12032 Minute viru
9	1902.2	94.2	4764	14	MU034253	X02453 Mouse parvo
10	1902.2	94.2	5087	14	PAMVMI	X02481 Mouse parvo
11	1899	94.1	5144	14	MPUI2469	U12469 Mouse parvo
12	1873	92.8	4764	14	MU034254	U34255 Hamster par
13	1862.2	92.2	4773	14	HOU34255	U34256 Mice minute
14	1855.8	91.9	4761	14	MVU34256	CQ786765 Sequence
15	1771	87.7	5121	6	CQ786765	AF331230 Sequence
16	1662.2	82.3	4904	14	AF331230	AF332884 Rat minut
17	1659	82.2	4795	14	AF332884	U79033 Kilham rat
18	1657.4	82.1	4927	14	KR079033	M81888 Parvovirus
19	1656.6	82.1	5135	14	PVRSEQ	

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARHI	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	52.0	5124	14	PVPFVC	M38246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCYIA	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAPFVNS1	X55115 Feline panl
40	1043.6	51.7	5124	14	PVPFVC	M38245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1040.4	51.5	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVRPFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137743	Sequence 8 from Patent EP1077260.	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137743					
DEFINITION	AX137743					
ACCESSION	AX137743.1	GI:14273917				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
DEUTSCHES Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)						

Mice minute virus

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

Nuesch, J. and Rommelaere, J.  
Parvovirus ns1 variants  
Patent: EP 1077260-A 8 21-FEB-2001;  
Deutsches Krebsforschungszentrum

Location/Qualifiers  
1. .2019  
/organism="Mice minute virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10794"  
/note="Parvovirus NS1 variant"  
1. 2019  
/notes="unassigned protein product"  
/codon\_start=1  
/protein\_id="CAC39993.1"  
/db\_xref="GI:14273918"  
/translation="MAGNAYSDEVLGATNWLKESNOEVFSFVKFNENVLNGKDIGW  
NSYKKELODELKSLQGEATTTDQSEDMETTTVDKTKQVFIKDSLVKKCLFEVL  
NTNIFPGVNNFVQHWKQDQGHCHVLIGKDFSAQGWQRQNLVYNSRWLVT  
CNVQTPPABRIKLREIAEDNWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKKIS  
TSPRRDGGYFLSDSGWNTFLKGERHLVSKLYTDMPETVETTTTAQETKGRIS  
QTKKEYSIKTKLEKLVKRVTSPEDEMMQPSYIEMMAQPGGKGLNLTCTLT  
ARTKAFDLILEKETSLLTSLPDRACRIFAFHGMVYKVCIAICCVLNRQGGKR  
NTVLEHGPASTGKSIIAQIAQAVGNVGNANVNFENDCTNKLIVVEEAGNFQ  
QVNFKAICSGQITRIDQKSGSKQIEPTVIMTNNTNIVRIGKEEPHEHTQPIRD  
RLNLTHTLTFGLDFGLNBNPFI CAWLXNGYOSTWASYKAGKGVDPWENNAE  
PKVPTPINLIGARSFTPTKSTPLSQNTALTPLASDLEDLALEPLSTNPVAGTAE  
TQNTGSAQKACQDQGLSPTWSIEEDLRACFGAELPKKDFSEPLND"

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

	Best Local Similarity	100.0%; Pred. No. 0;	Mismatches	Indels	Gaps	
	Conservative	0;	0;	0;	0;	
Qy	1	ATGGCTGGAATAAGTCTACTCTCATCAAGGTTCCTTTGGGAGCAACCAACTGGTTTAAGCGAAAAA	60			
Dd	1	ATGGCTGGAANTGCATTACTCTCATCAAGGTTCCTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	60			
Qy	61	AGTAAACCAGGAAGTGTTCTCATTGTTTTTAAAATGAAAATGTTCAAACCTGAATCGAAAA	120			
Dd	61	AGTAAACCAGGAAGTGTTCTCATTGTTTTTAAAATGAAAATGTTCAAACCTGAATCGAAAA	120			
Qy	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGGACGAGCTGGAATCTTTTACA	180			
Dd	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGGACGAGCTGGAATCTTTTACA	180			
Qy	181	CGAGGAGCGGAAAACTACTTTGGGCCAACAGCGAGCACATGGGAATCGGAAACCCACAGTGGAT	240			
Dd	181	CGAGGAGCGGAAAACTACTTTGGGCCAACAGCGAGCACATGGGAATCGGAAACCCACAGTGGAT	240			
Qy	241	GAATGACCAAAAAACGAAGTATTCATTTTTGATTCTTTGGTTAAAAAATGTTTATTTTCAA	300			
Dd	241	GAATGACCAAAAAACGAAGTATTCATTTTTGATTCTTTGGTTAAAAAATGTTTATTTTCAA	300			
Qy	301	GTGCTTAACACAAAGAATATATTTCTCGTGATGTTTAATTGGTTTGTGCACATGAATGG	360			
Dd	301	GTGCTTAACACAAAGAATATATTTCTCGTGATGTTTAATTGGTTTGTGCACATGAATGG	360			
Qy	361	GGAAAAAGCAAGCGTGCACCTGCATGTACTTAATTCGGAGAAAGGACTTTTAGTCAAGCT	420			
Dd	361	GGAAAAAGCAAGCGTGCACCTGCATGTACTTAATTCGGAGAAAGGACTTTTAGTCAAGCT	420			
Qy	421	CAAGGGAATGGTGGAGAACCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	480			
Dd	421	CAAGGGAATGGTGGAGAACCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	480			
Qy	481	TGTAATGTGCACCTAACACAGCTGAAAGAAATTAAACTTAAGAGAAATAGCAGAGAACAAT	540			
Dd	481	TGTAATGTGCACCTAACACAGCTGAAAGAAATTAAACTTAAGAGAAATAGCAGAGAACAAT	540			
Qy	541	GAGTGGGTACTCTACTTACTTATAAGCATANGCAAAACCCAAGGACTATACCAAGTGT	600			
Dd	541	GAGTGGGTACTCTACTTACTTATAAGCATANGCAAAACCCAAGGACTATACCAAGTGT	600			
Qy	601	GTTCTTTTGGAAAAACATGATGCTTACTATTTTTTAACTAAAAAGAAATAAGCAGCTAGT	660			
Dd	601	GTTCTTTTGGAAAAACATGATGCTTACTATTTTTTAACTAAAAAGAAATAAGCAGCTAGT	660			
Qy	661	CCAACAAGAGACGAGGCTATTTTCTTAGCAGTGAATCTCGTGCTGGAAAACTAACTTTTTA	720			
Dd	661	CCAACAAGAGACGAGGCTATTTTCTTAGCAGTGAATCTCGTGCTGGAAAACTAACTTTTTA	720			
Qy	721	AAGAAGCGCGAGCGCCACTAGTCAGCANAATAACACTGATGATGCTGCGGCAGAAACG	780			
Dd	721	AAGAAGCGCGAGCGCCACTAGTCAGCANAATAACACTGATGATGCTGCGGCAGAAACG	780			
Qy	781	GTTGAAACCAACAGTAAACCACTGCGCAGAAACTAAGCGCGCAGAAATTCAAACTAAAAA	840			
Dd	781	GTTGAAACCAACAGTAAACCACTGCGCAGAAACTAAGCGCGCAGAAATTCAAACTAAAAA	840			
Qy	841	GAAGTTTCTATTAAAACTTACAATTAAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	900			
Dd	841	GAAGTTTCTATTAAAACTTACAATTAAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	900			
Qy	901	GACTTGGATGATGTCAGCCNAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA	960			
Dd	901	GACTTGGATGATGTCAGCCNAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA	960			
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTGTACACTTAACCTTAGCCAGAACCAAAAAACAGCA	1020			
Dd	961	AACCTGCTGAAAAATACGCTAGAGATTGTACACTTAACCTTAGCCAGAACCAAAAAACAGCA	1020			
Qy	1021	TTTGACTTTAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC	1080			
Dd	1021	TTTGACTTTAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC	1080			

Db	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACACAGCAAACTAAACCAACTTTTCACTGCCTGAC	1081
Qy	1081	ACAAGAGCCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT	1140
Db	1081	ACAAGAGCCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT	1140
Qy	1141	ATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTATTTCATGACCA	1200
Db	1141	ATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTATTTCATGACCA	1200
Qy	1201	GCCAGCACAGGCAATCTATTATTGCAACAAGCCATAGCACAAAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCCAGCACAGGCAATCTATTATTGCAACAAGCCATAGCACAAAGCAGTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATGCAGCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAAGAACTTTGATT	1320
Db	1261	TGCTATAATGCAGCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAAGAACTTTGATT	1320
Qy	1321	TGGGTAGAAGAAGCTGCTAACTTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTTGCTCT	1380
Db	1321	TGGGTAGAAGAAGCTGCTAACTTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTTGCTCT	1380
Qy	1381	GGTCAAACTATTTCGCATTTGATTCAAAAGGAAAAGGCGCAAAACAGATTGAAACCAACCA	1440
Db	1381	GGTCAAACTATTTCGCATTTGATTCAAAAGGAAAAGGCGCAAAACAGATTGAAACCAACCA	1440
Qy	1441	GTCATCATGACCAAAATGAGAACATTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Db	1441	GTCATCATGACCAAAATGAGAACATTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Qy	1501	GAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCAATCAACACATACCTTGCCT	1560
Db	1501	GAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCAATCAACACATACCTTGCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGTGACAAAATAGAAATGCGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGTGACAAAATAGAAATGCGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
Qy	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGCTTAAATGGGGCAAGTTCTCGATTGCG	1680
Db	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGCTTAAATGGGGCAAGTTCTCGATTGCG	1680
Qy	1681	TCGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATATAATTTACTAGGTTTCGGCACGC	1740
Db	1681	TCGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATATAATTTACTAGGTTTCGGCACGC	1740
Qy	1741	TCACATTCACGACACCGGAAAAGTAGCGCTCTCAGCCAGAACTATGCGACTAATCTCACTT	1800
Db	1741	TCACATTCACGACACCGGAAAAGTAGCGCTCTCAGCCAGAACTATGCGACTAATCTCACTT	1800
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTCGAGCACACCAATACTCTCGTTGCG	1860
Db	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTCGAGCACACCAATACTCTCGTTGCG	1860
Qy	1861	GGCACTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAAGCCTGCCAAGATGCTCAA	1920
Db	1861	GGCACTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAAGCCTGCCAAGATGCTCAA	1920
Qy	1921	CTGAGCCCAACTTTGGTTCAGAGATTCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACCG	1980
Db	1921	CTGAGCCCAACTTTGGTTCAGAGATTCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACCG	1980
Qy	1981	TTGAAGAAAGACTTCAGCGGAGCGCTGAACTTGGACTAA	2019
Db	1981	TTGAAGAAAGACTTCAGCGGAGCGCTGAACTTGGACTAA	2019

RESULT 2

AX137736

LOCUS

AX137736

DEFINITION

Sequence 1 from Patent EP1077260.

ACCESSION

AX137736

VERSION

AX137736.1

GI:14273909

2019 bp

DNA

linear

PAT 30-MAY-2001

RESULT 2	2019 bp	DNA	linear	PAT 30-MAY-2001
AX137736				
LOCUS	AX137736	Sequence 1 from Patent EP1077260.		
DEFINITION	AX137736			
ACCESSION	AX137736.1	GI:14273909		
VERSION				

KEYWORDS	Mice minute virus	Db	481	TGTAATGTGCAACTAACACGAGCTGAAAGNATTAATACTAAGAGAAATAGCAGAGCAAT	540
SOURCE	Mice minute virus	Qy	541	GAGTGGTTACTCTACTTACTTATATAGCATAGCAAAACCAAAAGAAAGACTATACAAAGTGT	600
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	Db	541	GAGTGGTTACTCTACTTACTTATATAGCATAGCAAAACCAAAAGAAAGACTATACAAAGTGT	600
REFERENCE	1 Nueesch, J. and Rommelaere, J.	Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATAAGCACTAGT	660
AUTHORS	Parvovirus ns1 variants	Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATAAGCACTAGT	660
TITLE	Patent: EP 1077260-A 1 21-FEB-2001;	Qy	661	CCACCAAGAGACGAGGCTTATTTCTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)	Db	661	CCACCAAGAGACGAGGCTTATTTCTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
FEATURES	Location/Qualifiers	Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG	780
source	1. .2019	Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG	780
CDS	/organism="Mice minute virus"	Qy	781	GTGTAACACACAGTAACTGCGGAGGAACTTAAGCGGCGAGAAATTCAAACTTAAAAA	840
	/mcl_type="unassigned DNA"	Db	781	GTGTAACACACAGTAACTGCGGAGGAACTTAAGCGGCGAGAAATTCAAACTTAAAAA	840
	/db_xref="taxon:10794"	Qy	841	GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAAGAGTAACTCACCAGAG	900
	/note="Wildtype Parvovirus NS1"	Db	841	GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAAGAGTAACTCACCAGAG	900
	1. .2019	Qy	901	GACTGATGATGATGCGAGCAGACAGTTACATTTGAAATGATGGCTCAACAGGTTGGAGAA	960
	/note="unnamed protein product"	Db	901	GACTGATGATGATGCGAGCAGACAGTTACATTTGAAATGATGGCTCAACAGGTTGGAGAA	960
	/codon_start=1	Qy	961	AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAACAGCA	1020
	/protein_id="CAC39989.1"	Db	961	AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAACAGCA	1020
	/db_xref="GI:14273910"	Qy	1021	TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAAACCACTTTTCACTGCCTGAC	1080
	/translation="MAGNAYSDEVLGATNWLKEKSNQEVFSFVFNENVLQKDIGW	Db	1021	TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAAACCACTTTTCACTGCCTGAC	1080
	NSYKELQDELKSLQGAETTWDOSEDMEWETTVDEMTKOVFIQFDSLKKCLPEVL	Qy	1081	ACAAGAGCCTGCAAGATTTTTCCTTTTCATGGCTGGAACCTATGTTAAAGTTTCCCATGCT	1140
	NTKNIPFGDWNFVQHEWKDQGWCHVLIIGKDFSOAGKWRBQLNYYWRLVTA	Db	1081	ACAAGAACCTGCGAGAAATTTTCTTTTCATGGCTGGAACCTATGTTAAAGTTTCCCATGCT	1140
	CNVQLTPABRIKLREIAENENWTLTYKHQTKDYTKCVLFGNMAYFLTKKIS	Qy	1141	ATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTATGAGCA	1200
	TSPPRDGGYFLSSDSKDFLKEGERHLVSKSYIDMMAPQPGENLLKNTLEICTL	Db	1141	ATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTATGAGCA	1200
	QTKVEISIKTLKELVHKRVTSPEDMMQPSYIEMMAQPGENLLKNTLEICTL	Qy	1201	GCAGCAGAGGCAAACTCTATTATTGCACAAAGCCATAGCAGCAAGCAGTGGCAATGTTGT	1260
	ARTAFADILILEKABTSKLTNPSLPTDTRCIPAFHGMNVKCHAIICVLNRQGR	Db	1201	GCAGCAGAGGCAAACTCTATTATTGCACAAAGCCATAGCAGCAAGCAGTGGCAATGTTGT	1260
	NTVLFHPASTGKSIIAOIAOAGNVGYNAAVNFPENDCTNKNLIWVERAGNFG	Qy	1261	TGCTATAATGCGAGCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT	1320
	QYNPKALCSGOTIRIDOKGSKOIETPTVLTNNENITVVRIGCEEPHTOPIRD	Db	1261	TGCTATAATGCGAGCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT	1320
	RMNLHLHTLGRDGLVDXNHPMICALWLVKNGYQSTWASCAKNGKVPDSEWAE	Qy	1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTAAAGCACTTTTGCCTCT	1380
	PTPTPINILGSGSFPTPKSPUSQNTALPTLASDLEDLALPEWSTPTNPVAGTAE	Db	1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTAAAGCACTTTTGCCTCT	1380
	TQNTGEAGSKACQDQGLSPTWSEIEDLRACFAPLKKDFSEPLND"	Qy	1381	GGTCAAACTATTTCGGCATTCATCAAAAGGAAAGGCAAGCAAGATTGAACCAACCA	1440
ORIGIN		Db	1381	GGTCAAACTATTTCGGCATTCATCAAAAGGAAAGGCAAGCAAGATTGAACCAACCA	1440
Query Match	99.9%; Score 2017.4; DB 6; Length 2019;	Qy	1441	GTTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Best Local Similarity	100.0%; Pred. No. 0;	Db	1441	GTTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Matches 2018; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	Qy	1501	GAAACACATCAACCAATCAGACAGAAATGCTTTAACTTCACTTAACATACATCTTGCTCT	1560
1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	Db	1501	GAAACACATCAACCAATCAGACAGAAATGCTTTAACTTCACTTAACATACATCTTGCTCT	1560
1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	Qy	1561	GGTGTACTTTGGTTGGTGACAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
61	AGTAACGAGAGAGTGTCTCATTTGTTTAAATGAAAAATGTTCAACTGAATGAAAA	Db	1561	GGTGTACTTTGGTTGGTGACAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
61	AGTAACGAGAGAGTGTCTCATTTGTTTAAATGAAAAATGTTCAACTGAATGAAAA	Qy			
121	GATATCGGATGNAATGTTACAAAAAGAGCTGACGAGGAGCGAGCTGAAATCTTTACAA	Db			
121	GATATCGGATGNAATGTTACAAAAAGAGCTGACGAGGAGCGAGCTGAAATCTTTACAA	Qy			
181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	Db			
181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	Qy			
241	GAAATGACCAAAAGCAAGTATTCATTTTGGATTTCTGTTTAAATGTTTAAATGTTAA	Db			
241	GAAATGACCAAAAGCAAGTATTCATTTTGGATTTCTGTTTAAATGTTTAAATGTTAA	Qy			
301	GTGCTTTAACACAAAGAAATATTTCTGCTGATGTTTAAATGTTTGTGCAACATGAAATGG	Db			
301	GTGCTTTAACACAAAGAAATATTTCTGCTGATGTTTAAATGTTTGTGCAACATGAAATGG	Qy			
361	CGAAAGACCAAGGCTGGCACTGCCATGTACTTAATGGAGGAAAGGACCTTTAGTCAAGCT	Db			
361	CGAAAGACCAAGGCTGGCACTGCCATGTACTTAATGGAGGAAAGGACCTTTAGTCAAGCT	Qy			
421	CAAGGAAATGGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	Db			
421	CAAGGAAATGGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	Qy			
481	TGTAATGTGCAACTAACACGAGCTGAAAGNATTAATACTAAGAGAAATAGCAGAGCAAT	Db			

QY	1621	AATGGTTTACCAATCTACCATGGCAAGTACTGTGCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
	1621	AATGGTTTACCAATCTACCATGGCAAGTACTGTGCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
QY	1681	TCAGAAATCTGGCGGAGCAAAAGTGCCAACTCTTATATAATTTACTAGTTCGGCAGCG	1740
	1681	TCAGAAATCTGGCGGAGCAAAAGTGCCAACTCTTATATAATTTACTAGTTCGGCAGCG	1740
QY	1741	TCACCATTCAGCACCGAAAGTACGCCTCTCAGCAGAACTATGCACTAATCACTT	1800
	1741	TCACCATTCAGCACCGAAAGTACGCCTCTCAGCAGAACTATGCACTAATCACTT	1800
QY	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCTTTGAGCAGACACCAATATCTCTGTTGG	1860
	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCTTTGAGCAGACACCAATATCTCTGTTGG	1860
QY	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGCTCAA	1920
	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGCTCAA	1920
QY	1921	CTGAGCCCACTTGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG	1980
	1921	CTGAGCCCACTTGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG	1980
QY	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTCGACTAA	2019
	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTCGACTAA	2019
RESULT 3			
LOCUS	PAMV2	5081 bp	DNA linear VRL 10-FEB-1999
DEFINITION	Minute virus of mice with two major open reading frames (genome).		
ACCESSION	V01115		
VERSION	V01115.1	GI:60911	
KEYWORDS	coat protein; genome; origin of replication; overlapping genes; terminal repeat.		
SOURCE	Mice minute virus		
ORGANISM	Mice minute virus		
REFERENCE	1 (bases 1 to 5081)		
AUTHORS	Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.		
TITLE	The complete DNA sequence of minute virus of mice, an autonomous parvovirus		
JOURNAL	Nucleic Acids Res. 11 (4), 999-1018 (1983)		
MEDLINE	83143341		
PUBMED	6298737		
COMMENT	The messenger RNA of this virus (colinear with the strand listed) is spliced, but the exact splice sites are not known. The listed strand is complementary to the one which is included in the virion.		
FEATURES	Location/Qualifiers		
source	1..5081		
	/organism="Mice minute virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10794"		
	114..2279		
	/notes="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24309.1"		
	/db_xref="GI:60912"		
	/db_xref="GOA:P03134"		
	/db_xref="Swiss-Prot:P03134"		
	/translation="MISGSGSLNKGAKRWKFWKVVYKQLKSVYLVLFHVSVDQAKE SNLTWAGNAYSDVLGATNMLKEKSNQEVFVKNENVLQKDIWNYSVKELQED DELSLQRAETWDSSEMETHETVDETKQVIFPDSLVKCLFVLNLTKNIPFQD VNWFOHEWKGQDGHVLLIGKQFQSAQKWRRLNVLNWSRLVLTACNVQLTPAE RIKLEIAEDNEFWLLTYKHKQDYTKVLFGNMIAAYFLTKKKISTSPRGGY FLSSDGMKTNFLKEGERHLVSKLYDDMRPETVETVTAQETKRGRTQTKKYSIK TTLKELVHKRVTSPEWMMQPDVSEMAQPGENLLNTLEICTLTLARTKTAFLD ILEKETSKLTFNPSLPDTCRIEAPHGNVYKVCALICVLNRQGGKENTVLFHGA STGKSLIAQALQAQVNGVCYNAANVFFNDCTKNLILWSEAGNFGQOVQFKAIC SGQTIRIDQKQKQIEPTFVIMTNNENITVVRIGCEERPEHTQPIRDMNLHLTH		
CDS			
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24310.1"		
	/db_xref="GI:60913"		
	/db_xref="GOA:P03137"		
	/db_xref="Swiss-Prot:P03137"		
	/translation="MAPPRAKRAKRGKLRDGLWLVY"		
	2332..2361		
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24311.1"		
	/db_xref="GI:584078"		
	/db_xref="TrEMBL:O89491"		
	/translation="MVGWMGINV"		
	2354..2398		
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24312.1"		
	/db_xref="GI:60914"		
	/db_xref="TrEMBL:O89818"		
	/translation="MFNYLFYRPEITWF"		
	2383..4551		
	/note="unlabeled protein product; Protein sequence is in conflict with the conceptual translation; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24313.1"		
	/db_xref="GI:1335397"		
	/db_xref="GOA:O84207"		
	/db_xref="TrEMBL:O84207"		
	/translation="NHLVLGVWPPGKYKVLPGNSLDQCEPTNPSDAAKEHDEAYDQY IKSGKNPDIYSAADQRFIDQDKADWGGKRGHYFFRTKRAFAPKLAIDSEPGTSGV ADGPGGGGGGGGGVGTSGYDNOYHREFLGDGWVEITATRLVHLNPKSEY CRIRVHNTDTSVKGNMAKDAHQIWTWPSLVADANWGLQPSDWQVICTMTSGLN LVSQDEIPNVVLKTVTDQDLGQAIKLYNNDLTACMVAVDNNILPYTPAANSMET LGFPWKPETIASPVRYEYCDVDLSVTYENQEGTVENVMGTPEKGIPOFFTTTQOI TLIRTGDEFATGTYTTFDINSVKLITHTQNLQGPPLLSFPPEADTAGTILTAQGR HGTTQMGVNWSEAIRTPAQVGFCCPHNDFEASRAGFPAPKVPADITQGVDRKANG SVRYSYKGHGENWASHGPAERYETWDSFSGSDTKDGFQSAFLVVPPLNGILT NANPTGTNDIHFNSVNSYGLPTAFSPHSVPYQGIWDEKDELHKLPRHITAPFV CKNAPAGQMLVRLGPNLTDQYDPNGATLSRIYVTGTFWKGKLTWRKLANTTWNPV YQVSAEDNGNSVMSVTKLPTATGNWQSVPLITRPVARNY"		
ORIGIN			
	Query Match 99.9%; Score 2017.4; DB 14; Length 5081;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	ATGGCTGGAATGCTTACTCTGATCAAGTTTGGGAGCAACCACTGGTTAAGGAAAA	60
DB	261	ATGGCTGGAATGCTTACTCTGATCAAGTTTGGGAGCAACCACTGGTTAAGGAAAA	320
QY	61	AGTAACCAAGAGTGTCTCTATTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
DB	321	AGTAACCAAGAGTGTCTCTATTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	380
QY	121	GATATCGGATGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAATCTTTTACAA	180
DB	381	GATATCGGATGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAATCTTTTACAA	440
QY	181	CGAGAGCGGAACACTACTTGGACCAACCGAGGACATGGAATGGGAACACACAGTGGAT	240
DB	441	CGAGAGCGGAACACTACTTGGACCAACCGAGGACATGGAATGGGAACACACAGTGGAT	500
QY	241	GAAATGACCAAAAAACAAGTATTCTATTGATCTTTGGTTTAAAAAATGTTTATTGAA	300
DB	501	GAAATGACCAAAAAACAAGTATTCTATTGATCTTTGGTTTAAAAAATGTTTATTGAA	560
QY	301	GTGCTTAACAAAGAAATATATTTCTCGTGTGATGTAATTTGTTGTGCAACATGATCG	360

Db 561 GTGCTTACACAAAGAAATATATTTCTGTGATGTTAAATGGTTTGTGCAACATGAATGG 620  
QY 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGACATTTAGTCAAGCT 420  
Db 621 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGACATTTAGTCAAGCT 680  
QY 421 CAAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC 480  
Db 681 CAAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC 740  
QY 481 TGTAAATGTCACACTAACACAGCTGAAGAAATTAACACTAAGAAATAGCAGAAAGCAAT 540  
Db 741 TGTAAATGTCACACTAACACAGCTGAAGAAATTAACACTAAGAAATAGCAGAAAGCAAT 800  
QY 541 GAGTGGGTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT 600  
Db 801 GAGTGGGTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT 860  
QY 601 GTTCTTTTGGAAAACATGATGTTGTTACTTATTTTAACTTAAAGAAATAGCAGTGT 660  
Db 861 GTTCTTTTGGAAAACATGATGTTGTTACTTATTTTAACTTAAAGAAATAGCAGTGT 920  
QY 661 CCACCAAGAGCGGAGGCTATTTCTTACAGTGTACTCTGGCTGGGAAACTAACTTTTAA 720  
Db 921 CCACCAAGAGCGGAGGCTATTTCTTACAGTGTACTCTGGCTGGGAAACTAACTTTTAA 780  
QY 721 AAAGAGGCGGAGCGGCATCTAGTGAGCAAACTATACACTGATGACATGGCGGCAGAAAGC 780  
Db 981 AAAGAGGCGGAGCGGCATCTAGTGAGCAAACTATACACTGATGACATGGCGGCAGAAAGC 1040  
QY 781 GTTGAACACCAAGTAAACCACTGGCAGGAAACTAAGCGCGGAGAAATCAAACTAAAAAA 840  
Db 1041 GTTGAACACCAAGTAAACCACTGGCAGGAAACTAAGCGCGGAGAAATCAAACTAAAAAA 1100  
QY 841 GAAGTTCTTATTAACACTACACTTAAGAGCTGGTGTCATAAAGAGTACCTCACAGAG 900  
Db 1101 GAAGTTCTTATTAACACTACACTTAAGAGCTGGTGTCATAAAGAGTACCTCACAGAG 1160  
QY 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGGCTCAACAGAGTGAGAA 960  
Db 1161 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGGCTCAACAGAGTGAGAA 1220  
QY 961 AACCTGCTGAAAAAATACGCTAGAGATTTGTAACACTCTAGCCAGAACCAAAACAGCA 1020  
Db 1221 AACCTGCTGAAAAAATACGCTAGAGATTTGTAACACTCTAGCCAGAACCAAAACAGCA 1280  
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAACAGCAGCAAACTAACCACTTTTCACTGCTGAC 1080  
Db 1281 TTTGACTTAATTTAGAAAAAGCTGAACAGCAGCAAACTAACCACTTTTCACTGCTGAC 1340  
QY 1081 ACAAGAGCCTGCAGAAATTTTGTCTTTTATGCTGCGAACTATGTTAAAGTTTGCCATGCT 1140  
Db 1341 ACAGAACCTGCAGAAATTTTGTCTTTTATGCTGCGAACTATGTTAAAGTTTGCCATGCT 1400  
QY 1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTATTTATGAGCA 1200  
Db 1401 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTATTTATGAGCA 1460  
QY 1201 GCCAGACAGGCAAAATCTATTTATGCAAGGCCATAGCAACAGCAGTTGCGCAATGTTGT 1260  
Db 1461 GCCAGACAGGCAAAATCTATTTATGCAAGGCCATAGCAACAGCAGTTGCGCAATGTTGT 1520  
QY 1261 TGTATAATGACGCCAATGTAACCTTTTCACTTTAATGACTGTACCAACAGAACTTGATT 1320  
Db 1521 TGTATAATGACGCCAATGTAACCTTTTCACTTTAATGACTGTACCAACAGAACTTGATT 1580  
QY 1321 TGGGTAGAAAGAGCTGGTAACCTTTGACAGCAGTAAGCAAGTTTAAAGCCATTTGCTCT 1380  
Db 1581 TGGGTAGAAAGAGCTGGTAACCTTTGACAGCAGTAAGCAAGTTTAAAGCCATTTGCTCT 1640  
QY 1381 GGTCAAACTATTGCGCATGATCAAAAGGAAAGGAGGAGCAAAACAGATTGAACCAACACCA 1440

Db 1641 GGTCAAACTATTGCGCATGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACACCA 1700  
QY 1441 GTCATCATGACCAAAATGAGAACATTAACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500  
Db 1701 GTCATCATGACCAAAATGAGAACATTAACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1760  
QY 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACATTTCACTACATACATACCTTGGCCT 1560  
Db 1761 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACATTTCACTACATACATACCTTGGCCT 1820  
QY 1561 GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620  
Db 1821 GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1880  
QY 1621 AATGTTTACCAATCTACATGCAAGTACTGTGCTTAATGGGGCAAAAGTTCTCTGATTGG 1680  
Db 1881 AATGTTTACCAATCTACATGCAAGTACTGTGCTTAATGGGGCAAAAGTTCTCTGATTGG 1940  
QY 1681 TCAGAAACTGGGGAGGCGCAAGGTCCTCTATATAAATTTACTAGGTTGCGGACCG 1740  
Db 1941 TCAGAAACTGGGGAGGCGCAAGGTCCTCTATATAAATTTACTAGGTTGCGGACCG 2000  
QY 1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCCACTT 1800  
Db 2001 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCCACTT 2060  
QY 1801 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTGAGACACACCAAAATCTCTCTGTGG 1860  
Db 2061 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTGAGACACACCAAAATCTCTCTGTGG 2120  
QY 1861 GGCATGCGAAGAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGCTCAA 1920  
Db 2121 GGCATGCGAAGAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGCTCAA 2180  
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACCG 1980  
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACCG 2240  
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019  
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2279

## RESULT 4

MVMPG 5149 bp ss-DNA linear VRL 22-MAY-1995  
LOCUS Minute virus of mice, complete genome.  
DEFINITION J02275 M12520 M12521 M14704  
ACCESSION J02275.1 GI:332293  
VERSION alternative splicing; capsid protein; complete genome;  
KEYWORDS nonstructural protein.  
SOURCE Mice minute virus  
ORGANISM Mice minute virus  
REFERENCE 1 (bases 1 to 5149)  
AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.  
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus  
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)  
MEDLINE 83143341  
PUBMED 6298737  
REFERENCE 2 (bases 1 to 5149)  
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.  
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain  
JOURNAL J. Virol. 57 (2), 656-669 (1986)  
MEDLINE 86115415  
PUBMED 3502703  
REFERENCE 3 (sites)  
AUTHORS Morgan,W.R. and Ward,D.C.  
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs



CDS	2354..2398	Db	861	GTTC	TTTTTGGAAACATGATTGCTTACTATTTT	920
/gene="VP"						
/note="ORF2; putative"						
/codon_start=1						
/product="unknown protein"						
/protein_id="AAA67113.1"						
/db_xref="GI:825484"						
/translation="MFNYLFPYRPEITWF"						
2399..54557						
/gene="VP1"						
/number=2						
2794..4557						
/gene="VP1"						
/note="VP2"						
/codon_start=1						
/protein_id="AAA67114.1"						
/db_xref="GI:332296"						
/translation="MSDGTSDPSGNVAHSAARVERAADPGSGSGSGGGVGVST						
GSYDQTHYRFLFDGWVEITATRLVHLMPKSENYCRIRVHTTDTISVKGNMAKDD						
AHEQIWPMSLVADANAGVWLQSDWQYICNTMSQJNLVSLDOEIFNVVLKTVTSQDL						
GGQAIKIYNDLUTACMVAVDSNNILPYTPAANSMETLGFYPWKFTIASPYRYPCVD						
RDLSTVYENQEGTEVHNVTGPKMNSQSFTEENTQOITLLRTGDFEATGTYFYFDTS						
VKLHTWOTNRLOGQPLLSTPEADITAGLTIAQGRHGTTCMGVNWVSEAIRTRPA						
QVGFQCPHNDFFASRAGPPAPKVPADITQGVDKVKEANGSVRYSYKQKHGHNWASGPA						
PRYTWDSTSFSGRDTKDFIQSAPLVVPPPLNGILLTNAPIGTKNDLHFNVSNSY						
GLPTAFSPSPVYPOQIWDKELDLHKPLRLHITAPFVCKNNAPGGQMLVRLGPNLTDQ						
Query Match 99.9%; Score 2017.4; DB 14; Length 5149;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;						
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60			
Db	261	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	320			
QY	61	AGTAACCGAGAGTGTTCTCATTTGTTTTTAAATGMAATGTTCAACTGAATGAAAA	120			
Db	321	AGTAACCGAGAGTGTTCTCATTTGTTTTTAAATGMAATGTTCAACTGAATGAAAA	380			
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTTACAA	180			
Db	381	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTTACAA	440			
QY	181	CGAGGAGCGGAAACTACTCTGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	240			
Db	441	CGAGGAGCGGAAACTACTCTGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	500			
QY	241	GAAATGACCAAAAGCAAGTATTCATTTTGATTTCTTTGGTTAAAAAATGTTTATTGAA	300			
Db	501	GAAATGACCAAAAGCAAGTATTCATTTTGATTTCTTTGGTTAAAAAATGTTTATTGAA	560			
QY	301	GTGCTTTAACACAAAGAAATATTTCTCGTGATGTTTAAATGGTTTGTGCAACATGAATGG	360			
Db	561	GTGCTTTAACACAAAGAAATATTTCTCGTGATGTTTAAATGGTTTGTGCAACATGAATGG	620			
QY	361	GGAAAGACCAAGGCTGGCAGCTGCCATGTAATAATTTGGAGGAAGCACTTTAGTCAAGCT	420			
Db	621	GGAAAGACCAAGGCTGGCAGCTGCCATGTAATAATTTGGAGGAAGCACTTTAGTCAAGCT	680			
QY	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC	480			
Db	681	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC	740			
QY	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAAGTAAAGAGAAATAGCAGAGCAAT	540			
Db	741	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAAGTAAAGAGAAATAGCAGAGCAAT	800			
QY	541	GAGTGGGTTACTCTACTTACTTATAAGCATTAAGCAACCAAAAAAGACTATACCAAGTGT	600			
Db	801	GAGTGGGTTACTCTACTTACTTATAAGCATTAAGCAACCAAAAAAGACTATACCAAGTGT	860			
QY	601	GTTC	TTTTTGGAAACATGATTGCTTACTATTTT	TAACTAAAAAGAAATAAGCACTAGT	660	



```
Qy 1741 TCACCAATTCAGCACCGAAAGATACGCTCTCAGCCAGCAATATGCACTAACTCCACTT 1800
Db 2001 TCACCAATTCAGCACCGAAAGATACGCTCTCAGCCAGCAATATGCACTAACTCCACTT 2060
Qy 1801 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATATCTCTGTTGCG 1860
Db 2061 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATATCTCTGTTGCG 2120
Qy 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTCCCAAGATGGTCAA 1920
Db 2121 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTCCCAAGATGGTCAA 2180
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGGAACCG 1980
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGGAACCG 2240
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGAGCTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGAGCTAA 2279

RESULT 5
AX137739
LOCUS
DEFINITION
Sequence 4 from Patent EPI077260.
ACCESSION
AX137739
VERSION
AX137739.1 GI:14273913
KEYWORDS
Mice minute virus
SOURCE
ORGANISM
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1. Nueesch, J. and Rommelaere, J.
AUTHORS
Parvovirus ns1 variants
TITLE
Patent: EP 107260-A 4 21-FEB-2001;
JOURNAL
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)
FEATURES
source
Location/Qualifiers
1. .2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1. .2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273914"
/translation="MAGNAYSDEVLTGATNLKESNOEVFSFVKENVQLNKDITGW
NSYKELQDELKSLQGAETTWQSEDMEWETTVDEMTKOVFI PDSLKKLEPEVL
NTKNI PPGDVNMFVQHEWKGDOGHCHVLIGKDFSOAQGWRLQNLVYWSRLVTA
CNVQLTPAERIKLREIAEDNEWTLTYKHQTKDYTKCVLFGNMIAYFLTKKIS
TSPRGGFLSDSGKWNFLKEGERHLVSKLTDMPRETEVITVTAQETKKRRI
QTKKEVAIKTLKELVHKVTSPEMMWQPDYSIBMAOPGGENLLKNTLRICTL
ARTKFAFLILEKETSLSLPTDTRCRI FAFHGMNVKVCNLI CCVLNRQGGFR
NTVLPFGPASTGKSI IAOAIAOAVNGVNCYNAVNFPENDCNKLI WVEAGNFGQ
OVNPKAI CSQGTIRIDOKGSGKQIEPTVIMTNNENITVVRIGCEERPEHTPIRD
RMLNIHLTLPDGLSLVDKNWPMICAMLVKNGVQSTWASCAKWKGVDPNSWAE
PKVPTINILGARSFPTPKSTPLSQNTALPLASDLDELALPEMPTNTPVAGTAE
TQNTGAGSKACQDQGLSPTWSEIEEDLRACFGAEPLKKDFSEPLNLD"

Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTITTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 1 ATGGCTGGAATGCTTACTCTGATGAAGTITTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Qy 61 AGTAAACGAGGAAGTGTCTCAITTTGTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 1141 ATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTTATTCATGGACCA 1200
```



QY	1201	GCAGCAGGCAAAATCTATTATTCACAAAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
Db	1201		
QY	1261	GCAGCAGGCAAAATCTATTATTCACAAAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
Db	1261		
QY	1261	TGCTATAATGCAGCCCAATGTAAACTTTCCATTTTAATGACTGTACCAACAGAACTTGATT	1320
Db	1261		
QY	1321	TGCTATAATGCAGCCCAATGTAAACTTTCCATTTTAATGACTGTACCAACAGAACTTGATT	1320
Db	1321		
QY	1321	TGGGTAGAAGAGCTGGTAACTTTTGACAGCAGTAAGTAAACAGTTTAAAGCAATTTGCTCT	1380
Db	1321		
QY	1381	GGTCAAACTATTGCGATTTGATCAAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA	1440
Db	1381		
QY	1441	GGTCAAACTATTGCGATTTGATCAAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA	1440
Db	1441		
QY	1441	GTCATCATGACCAAAATGAGACATTACAGTGGTCAAGTAAAGTGGTGGCAAGAAAGACCA	1500
Db	1441		
QY	1501	GTCATCATGACCAAAATGAGACATTACAGTGGTCAAGTAAAGTGGTGGCAAGAAAGACCA	1500
Db	1501		
QY	1501	GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTGCCT	1560
Db	1501		
QY	1561	GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG	1620
Db	1561		
QY	1621	AATGGTTACCAATCTACCAATGGAACATTTACAGTGGTCAAGTAAAGTGGTGGTAAAG	1680
Db	1621		
QY	1681	TCAGAAAACTGGCGGAGCCAAAGTGGCAACTCTTAAATTTACTAGTTTCGGCAGCG	1740
Db	1681		
QY	1741	TCAGAAAACTGGCGGAGCCAAAGTGGCAACTCTTAAATTTACTAGTTTCGGCAGCG	1740
Db	1741		
QY	1741	TCACCAATTCAACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTCCACTT	1800
Db	1741		
QY	1801	GCATCGATCTCAGGACCTGGCTTTAGAGCTTTAGAGCTTTGAGACACCAATATCTCTGTCG	1860
Db	1801		
QY	1861	GGCACTGCAGAAACCCAGAACACTGGGAGCTGGTTCCAAAGCCCTGCCAAGATGTCAT	1920
Db	1861		
QY	1921	CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTGTCTCGGTGCGGAACCG	1980
Db	1921		
QY	1981	TTGAAGAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019
Db	1981		

RESULT 6  
AX137747  
LOCUS  
DEFINITION  
Sequence 12 from Patent EP1077260.  
ACCESSION  
AX137747  
VERSION  
AX137747.1 GI:14273921  
KEYWORDS  
Mice minute virus  
SOURCE  
Mice minute virus  
ORGANISM  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
REFERENCE  
1  
Nuesch, J. and Rommelaere, J.  
Parvovirus ns1 variants  
Patent: EP 1077260-A 12 21-FEB-2001;  
JOURNAL  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

(DE)

FEATURES  
source

Location/Qualifiers  
1. .2019  
/organism="Mice minute virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10794"  
/note="Parvovirus NS1 variant"  
1. .2019  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC39995.1"  
/db\_xref="GI:14273921"  
/translation="MAGNAYSDEVLTGATNWLKESNOEVESFVKENVQLNGKDLGW  
NSYKELQDELKSLQAGETTTWDSQEDNEWETTVDEMTKKQVIFDLSLVKKCLFEVL  
NTKNIPFGDNVFWQHEWQDQWHCHVLIIGGDFSDQAGKMWRLQNLVYWSRWLVIA  
CNVQITPAERIKLREIAEDNEWTLTYKHQTKYDTRKCVLFGNMIAFYFLTKKXIS  
TSPDDGGYFLSSDGWKTNFKLEGERHLVSKYTDMDRPEVETVTTTAAQETKGR1  
QTKKVS1KTLTKELVHKRVTSPEWMMQPDYSY1EMMAQPGGKLNKLTLEICTLTL  
ARTKAFDLILEKAETSKLTNLSPLDTRTCR1FAFHGWNVVKVCHAI CVDLNRQSKR  
NAVLFHPASTGKSTIAQAQVAGNVGVCYNANVNPENDCTNKNLIWVEERAGNGQ  
QVNFKAICSGQTRIDQKGGSKQIEPTVPTVIMTNENITVVRIGCEERPEHTQPIRD  
RMLNTHLTLPDFGLVDKNEWPMI CAWLKNGYQSTWASCAKWKVPDMSNNWAE  
PKVPTPINLLGSARGPFTPKPTPLSQNYALTPLASDLBDLALEPMSPTNPVAGTAE  
TQNTGEAGSKACQDGLSPTWSEIBEDLRACFGAEPLKKDFSEPLND"

ORIGIN

Query Match	99.8%;	Score	2015.8;	DB	6;	Length	2019;
Best Local Similarity	99.9%;	Pred. No.	0;				
Matches	2017;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60				
Db	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60				
QY	61	AGTAACCAAGAGTGTCTCATTTGTTTAAAAATGAAATCTTCAACTGAATGAAAA	120				
Db	61	AGTAACCAAGAGTGTCTCATTTGTTTAAAAATGAAATCTTCAACTGAATGAAAA	120				
QY	121	GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180				
Db	121	GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180				
QY	181	CGAGGAGCGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT	240				
Db	181	CGAGGAGCGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT	240				
QY	241	GAAATGACCAAAAGCAAGTATTTCAATTTGTTTCTTGTAAAAATCTTTTATTGAA	300				
Db	241	GAAATGACCAAAAGCAAGTATTTCAATTTGTTTCTTGTAAAAATCTTTTATTGAA	300				
QY	301	GTGCTTAAACAAAGAAATATATTTCTGTTGATGTTTAAATTTGTTTGTGCAACATGAATGG	360				
Db	301	GTGCTTAAACAAAGAAATATATTTCTGTTGATGTTTAAATTTGTTTGTGCAACATGAATGG	360				
QY	361	GGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420				
Db	361	GGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420				
QY	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC	480				
Db	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC	480				
QY	481	TGTAATGTGCAACTTAACACAGCTGAAGAAATTAAGAGAAATAGCAGAGACAAT	540				
Db	481	TGTAATGTGCAACTTAACACAGCTGAAGAAATTAAGAGAAATAGCAGAGACAAT	540				
QY	541	GAGTGGGTTACTTACTTATAAGCATTAAGCAAAACCAAAAAAGACTATACCAAGTGT	600				
Db	541	GAGTGGGTTACTTACTTATAAGCATTAAGCAAAACCAAAAAAGACTATACCAAGTGT	600				
QY	601	GTTCTTTTGGAAACATGATGCTGCTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660				
Db	601	GTTCTTTTGGAAACATGATGCTGCTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660				

661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720  
Db CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720  
721 AAAGAAGGCGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGCGCCAGAAACG 780  
Db AAAGAAGGCGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGCGCCAGAAACG 780  
781 GTTGAACACACAGTAACCTAGCGGAGGAACTTAAGCGGCGAGAACTTCAAACTAAAAAA 840  
Db GTTGAACACACAGTAACCTAGCGGAGGAACTTAAGCGGCGAGAACTTCAAACTAAAAAA 840  
841 GAAGTTTCTATTAAAACTACACTTAAGAGCTGGTGCCATAAAAGAGTAACCTCACACAG 900  
Db GAAGTTTCTATTAAAACTACACTTAAGAGCTGGTGCCATAAAAGAGTAACCTCACACAG 900  
901 GACTGATGATGATGACGAGCAGTGTACATTGAATGATGGCTCAACAGAGTGAGAA 960  
Db GACTGATGATGATGACGAGCAGTGTACATTGAATGATGGCTCAACAGAGTGAGAA 960  
961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020  
Db AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020  
1021 TTGACTTAATTTAGAAAAAGCTGAAACCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080  
Db TTGACTTAATTTAGAAAAAGCTGAAACCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080  
1081 ACAGAGCCTGCGAAATTTTGTCTTTTCATGGCTGGAATGATTTAAAGTTTGCATGCT 1140  
Db ACAGAGCCTGCGAAATTTTGTCTTTTCATGGCTGGAATGATTTAAAGTTTGCATGCT 1140  
1141 ATTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACCTGTTTATTTCATGACCA 1200  
Db ATTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACCTGTTTATTTCATGACCA 1200  
1201 GCAGACAGGCAAACTTATTATGCAAGCCATAGCAAGCAGTGTGGCAATGTTGCT 1260  
Db GCAGACAGGCAAACTTATTATGCAAGCCATAGCAAGCAGTGTGGCAATGTTGCT 1260  
1261 TGTATATAGCAGCAATGTAACTTTTCACTTAATGACTGTACCAACAGAACTTGATT 1320  
Db TGTATATAGCAGCAATGTAACTTTTCACTTAATGACTGTACCAACAGAACTTGATT 1320  
1321 TGGGTAGAGAGCTGCTGTAACCTTTGACAGCAAGTAAACAGTTTAAAGCATTTCCTCT 1380  
Db TGGGTAGAGAGCTGCTGTAACCTTTGACAGCAAGTAAACAGTTTAAAGCATTTCCTCT 1380  
1381 GGTCAAACCTATTGCAATTGATCAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440  
Db GGTCAAACCTATTGCAATTGATCAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440  
1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGTATGGCTGCGAAGAAAGCA 1500  
Db GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGTATGGCTGCGAAGAAAGCA 1500  
1501 GAACACACTCAACCAATCAGACAGCAAGTCTTAACATTATCATCAACATACCTTGCT 1560  
Db GAACACACTCAACCAATCAGACAGCAAGTCTTAACATTATCATCAACATACCTTGCT 1560  
1561 GGTGACTTTTGGTTGTTGACAAAAATGAATGGCCCATGATTGTGCTTTGGTTGTTAAG 1620  
Db GGTGACTTTTGGTTGTTGACAAAAATGAATGGCCCATGATTGTGCTTTGGTTGTTAAG 1620  
1621 AATGGTTACCAATCTACATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTTCTGATTGG 1680  
Db AATGGTTACCAATCTACATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTTCTGATTGG 1680  
1681 TCAGAAAACTGGCGGAGCAAGGTCCTTATTAATTTTACTAGTTTCGGGACCG 1740  
Db TCAGAAAACTGGCGGAGCAAGGTCCTTATTAATTTTACTAGTTTCGGGACCG 1740

1741 TCACCAATTCAGACACCAGAAAGTAGTCCCTCTCAGCAGAACTATGCACTAATCCACTT 1800  
Db TCACCAATTCAGACACCAGAAAGTAGTCCCTCTCAGCAGAACTATGCACTAATCCACTT 1800  
1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCCTGTTGCG 1860  
Db GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCCTGTTGCG 1860  
1861 GGCACCTGCAAGAAACCCAGAACACTGGGGAGGCTGGTTCCAAAGCCTGCCAAAGTGTCAA 1920  
Db GGCACCTGCAAGAAACCCAGAACACTGGGGAGGCTGGTTCCAAAGCCTGCCAAAGTGTCAA 1920  
1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTCGTTCGCGGAAACCG 1980  
Db CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTCGTTCGCGGAAACCG 1980  
1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 7  
AX137751  
LOCUS AX137751 2019 bp DNA linear PAT 30-MAY-2001  
DEFINITION Sequence 16 from Patent EP1077260.  
ACCESSION AX137751  
VERSION AX137751.1 GI:14273925  
KEYWORDS  
SOURCE  
ORGANISM  
Mice minute virus  
Mice minute virus  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
REFERENCE  
1  
AUTHORS Nueesch,J. and Rommelaere,J.  
TITLE Parvovirus ns1 variants  
JOURNAL Patent: EP 1077260-A 16 21-FEB-2001;  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)

FEATURES  
source  
1. .2019  
/organism="Mice minute virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10794"  
/note="Parvovirus NS1 variant"  
1. .2019  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="C39997.1"  
/db\_xref="GI:14273926"  
/translation="MAGNAYSDEVLGATNWLKEKNQVSPFVFNENVLNGKDIGW  
NSYKLEDELKSLQGAETTWQSEDEMETTVDKTKQVIFDSLKKCLFEVL  
NTKNTFPQDVNMFQHEWQKQCHVLIGGDFSAQKWRRLQVLYWRWLVA  
CNVOLTPAERIKLRIADENWVLLTYKHOTKKDYTKCVLFGNMIAYFLTKKIS  
TSPRDGGVFLSSDSGKTNFLKEGERHLVSKLTDMDPEPTEVTTTAOETKRGRI  
QTKEVSVITLKLKHVKVTSPEWMMQPDSEYEMMAQPGENLLKNTLEICTLTL  
AKTFAFDLILEKASTKLTNPSLPTRTCRIFAFHGMNVKVCCHAICCVLNKQGRK  
NTVLFHSPASTKSIIAQIAQAVNGVCYNAANVPFNDCNTKNKLIWVEEAGNFGQ  
QMNQKASGOAIRDKDGKSKQIEPTPVITMTNENITVVRIGCEERPEHTQIRID  
RMLNTHLTHTLPGDLVDKNEWPMICAWLVKNYGQSTWASCAKWKVPDMSENWAE  
PKVPTPINLGSARSPFTPKSTPLSQNYALTPLASDLDELALPEWSTPNTPVAGTAE  
TQNTGAGSKACQDQGLSPTWSEIBEDLRACFGAEPLKKDFSEPLND"

CDS  
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN  
1 ATGGCTGGAATGCTTACTCTGATGAACTTTTGGGAGCAACCAACTGGTTAAGGAAAA 60  
Db ATGGCTGGAATGCTTACTCTGATGAACTTTTGGGAGCAACCAACTGGTTAAGGAAAA 60  
61 AGTACCAAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
Db AGTACCAAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120

121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA 180  
121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA 180  
181 CGAGGAGCGGAAACTACTTCGGACCAAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 240  
181 CGAGGAGCGGAAACTACTTCGGACCAAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 240  
241 GAAATGACCAAAAAAGCAAGTAATTCATTTTGAATCTTTTGGTTTAAAAAATTTTATTGAA 300  
241 GAAATGACCAAAAAAGCAAGTAATTCATTTTGAATCTTTTGGTTTAAAAAATTTTATTGAA 300  
301 GTGCTTAAACAAAGAAATATATTTCTGCTGGTGAATTTAAATTTGGTTTGTGCAACATGAATGG 360  
301 GTGCTTAAACAAAGAAATATATTTCTGCTGGTGAATTTAAATTTGGTTTGTGCAACATGAATGG 360  
361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTGCAAGCT 420  
361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTGCAAGCT 420  
421 CAAGGGAATTTGGTGGAGAGGCAACTAAATTTTACTTGGAGCAGATGGTTGTAACAGCC 480  
421 CAAGGGAATTTGGTGGAGAGGCAACTAAATTTTACTTGGAGCAGATGGTTGTAACAGCC 480  
481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 540  
481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 540  
541 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAGAGCACTATACCAAGTGT 600  
541 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAGAGCACTATACCAAGTGT 600  
601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT 660  
601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT 660  
661 CACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720  
661 CACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720  
721 AAGAAGGCGAGCGCCATCTAGTGAGCAACTATACACTGATCAGATGCGGCCAGAAACG 780  
721 AAGAAGGCGAGCGCCATCTAGTGAGCAACTATACACTGATCAGATGCGGCCAGAAACG 780  
781 GTTGAACCAACAGTAACCACTGCGCAGGAACTAAGCGCGGCAAGATTCAAACTAAAAA 840  
781 GTTGAACCAACAGTAACCACTGCGCAGGAACTAAGCGCGGCAAGATTCAAACTAAAAA 840  
841 GAAGTTTCTATTAAAACTACATTAAGAGCTGGTGCAAAAAAGAGTAACCTCACAGAG 900  
841 GAAGTTTCTATTAAAACTACATTAAGAGCTGGTGCAAAAAAGAGTAACCTCACAGAG 900  
901 GACTGGATGATGATGAGCAGACAGTTACATTTGAATGATGCTCAACAGAGTGGAGAA 960  
901 GACTGGATGATGATGAGCAGACAGTTACATTTGAATGATGCTCAACAGAGTGGAGAA 960  
961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
1021 TTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC 1080  
1021 TTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC 1080  
1081 ACNAGAGCTGCAAGATTTTGTCTTTTTCATGCTGGAATGATGCTTAAAGTTTCCATGCT 1140  
1081 ACNAGAGCTGCAAGATTTTGTCTTTTTCATGCTGGAATGATGCTTAAAGTTTCCATGCT 1140  
1141 ATTTGCTGTGTTTTAAACAGACAGGAGGCAAGAAATACTGTTTTTATTTTATGACCA 1200  
1141 ATTTGCTGTGTTTTAAACAGACAGGAGGCAAGAAATACTGTTTTTATTTTATGACCA 1200

1201 GCAGACAGGCAAAATCTATTATTGCAAGCCATAGCACAGCAGTTGGCAATGTTGGT 1260  
1201 GCAGACAGGCAAAATCTATTATTGCAAGCCATAGCACAGCAGTTGGCAATGTTGGT 1260  
1261 TGTCTAATGCAAGCAATGTAAACTTTTCCATTTTAACTGTACCAACAAGAACTTGATT 1320  
1261 TGTCTAATGCAAGCAATGTAAACTTTTCCATTTTAACTGTACCAACAAGAACTTGATT 1320  
1321 TGGGTAGAAAGACTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
1321 TGGGTAGAAAGACTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCAACCA 1440  
1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCAACCA 1440  
1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGCTGCGAAGAAAGACCA 1500  
1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGCTGCGAAGAAAGACCA 1500  
1501 GAAACACATCAACCAATTCAGAGACAGAACTTTAAACATTCATCTAACACATACCTTGCT 1560  
1501 GAAACACATCAACCAATTCAGAGACAGAACTTTAAACATTCATCTAACACATACCTTGCT 1560  
1561 GGTGACTTTGGTTGGTTGACAAAAAATGAATGSCCCCATGATTTGTGCTTGGTAAAG 1620  
1561 GGTGACTTTGGTTGGTTGACAAAAAATGAATGSCCCCATGATTTGTGCTTGGTAAAG 1620  
1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAAAGTTTCTGATTGG 1680  
1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAAAGTTTCTGATTGG 1680  
1681 TCAGAAAACTGGCGGAGCCAAAGTGCCTTAACTTAAATTTACTAGTTTGGCAGCG 1740  
1681 TCAGAAAACTGGCGGAGCCAAAGTGCCTTAACTTAAATTTACTAGTTTGGCAGCG 1740  
1741 TCACCAATTCAGACACCGAAAAAGTAGTCCCTCTCAGCAGAACTATGCACTAACCTCACTT 1800  
1741 TCACCAATTCAGACACCGAAAAAGTAGTCCCTCTCAGCAGAACTATGCACTAACCTCACTT 1800  
1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG 1860  
1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG 1860  
1861 GGCATCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA 1920  
1861 GGCATCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA 1920  
1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980  
1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980  
1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

## RESULT 8

MMICG 5085 bp ss-DNA linear VRL 05-JUL-1995  
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,  
DEFINITION complete genome.  
ACCESSION M12032  
VERSION M12032.1 GI:332289  
KEYWORDS alternative splicing; capsid protein; complete genome;  
nonstructural protein.  
SOURCE Mice minute virus  
ORGANISM Mice minute virus  
REFERENCE 1 (bases 1 to 5085)  
AUTHORS Åstell,C.R., Gardiner,E.M. and Tattersall,P.  
TITLE DNA sequence of the lymphotropic variant of minute virus of mice,  
MVM(i), and comparison with the DNA sequence of the fibrotropic

JOURNAL MEDLINE COMMENT	prototype strain J. Virol. 570, 656-669 (1986) 86115415 Original source text: Minute virus of mice (lymphotropic variant of MVM) DNA, clone pBG222. Location/Qualifiers 1. .5085 /organism="Mice minute virus" /mol_type="genomic DNA" /isolate="lymphotropic variant" /db_xref="taxon:10794" /clone="pBG222" 115. .2280 /gene="NS1" 115. .2280 /gene="NS1" /note="putative" /codon_start=1 /product="nonstructural protein" /protein_id="AAA69566.1" /db_xref="GI:825477" /translation="MISGESLIDQAKRKNMFKYVKOMLKSVTYPFPHSVSRDAQKESNQLTMAGNAYSDEVLTGTTNWLKESNQEVFVFVKTEDVQLNGKIDGNNYKKELQEDLKSLQGAETTWQSEDMEWESTVDEMTKKQVFIYDSLKKCLFEVLSTKNLIAPADVTWFVQHEGKQGWCHVLIIGKDFSQAGKWRRLNLYMSRWLTACNVQLTPAE RKLREIARDSEWTLITVKKTKDYTKCVLFGNMIAYFLTKKISTSPRDGGY FLSSDGWKTNFLKEGERHLVSKLYTDDMRPETVTTTAQETKRGRIOTKKEYSIK TILKELVHRVTSPEDMWMMQPDYSIEMMAQPGGENLLKNTLEICTLILARTKTAFDL ILEKATSKLTFNPSLPTDKIFAFPHGWNVYKVCICAICCVLNRRQGGKRNVTULFGPA STGKIIIAQIAQAVGNCYNAANVPFNCTNKNLIWVEEAGNFGQGVNQFKAIC SQGTIRIDQKGSQKIEPTVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTH TLPGDFGLVDKNEWPMICAWLVKNGYSTMASYCAKWKGVDPWSNWAEPKVPITINS LGSASPFPTPKSTPLSONYAITPLASDLELALPEWSTPNTPVAGTAETQNTGEAGS KACDQQLSPTWSEIEEDLRACGAEPKRDISEPLNLD" 201. .>2280 /gene="NS1" /product="NS-1 mRNA" 262. .2280 /gene="NS1" /codon_start=1 /product="nonstructural protein" /protein_id="AAA69567.1" /db_xref="GI:332291" /translation="MAGNAYSDEVLTGTTNWLKESNQEVFVFVKTEDVQLNGKIDGIM NNYKELQEDLKSLQGAETTWQSEDMEWESTVDEMTKKQVFIYDSLKKCLFEVL STKNLIAPADVTWFVQHEGKQGWCHVLIIGKDFSQAGKWRRLNLYMSRWLTAVT CNVQLTPABRIKLREIARDSEWTLITVKKTKDYTKCVLFGNMIAYFLTKKIS TSPPRDGGYFLSSDGSWKTNFLKEGERHLVSKLYTDDMRPETVTTTAQETKRGRIO TKKEYSIKTILKELVHRVTSPEDMWMMQPDYSIEMMAQPGGENLLKNTLEICTLIL ARTKTAFDLILEKATSKLTFNPSLPTDKIFAFPHGWNVYKVCICAICCVLNRRQGGK NTVLFHPGASTGKIIIAQIAQAVGNCYNAANVPFNCTNKNLIWVEEAGNFGQ OVNQPKAICSGOTIRIDQKGSQKIEPTVIMTNNITVVRIGCEERPEHTQPIRD RMLNIHLTHLFDGFLVDKNEWPMICAWLVKNGYSTMASYCAKWKGVDPWSNWA PKVPTPINSLGSASPFPTPKSTPLSONYAITPLASDLELALPEWSTPNTPVAGTAE TQNTGAGSKACDQQLSPTWSEIEEDLRACGAEPKRDISEPLNLD" 2003. .4558 /gene="VP" 2003. .2281 /gene="VP" /note="major transcription start site" /number=1 2007. .2281 /gene="VP" /note="minor transcription start site" /number=1 2010. .2281 /gene="VP" /note="minor transcription start site" /number=1 2282. .2399 /gene="VP" /note="alternative intron" 2282. .2377	gene CDS	/gene="VP" /note="alternative intron" join(2287. .2317,2400. .4558) /gene="VP1" join(2287. .2317,2400. .4558) /gene="VP1" /codon_start=1 /protein_id="AAA69569.1" /db_xref="GI:332290" /translation="MAPPKRAKRGWVPPGYKYLPGNSLDQGEPTNPSDAAAKEHDE AYDQIKSGKNPILYFSAADORFIDOTKADKDGKGVGHYFFRTKRAFAPKLATDSEP GTGVSRAKCKTRTPAYIFINQARAKKLTSSAAQQSQSTMSDGTSDQDGGNAVHSA RVERAADPGGGGGGGGGVGTSGYSDNQTHRYFLGDGWEVETALATRLVHLNMP KSENYCRIRVHNTTDSYKGNMAKDDAHEQIKNNDLTACMWAVDNNILPYTPAA MSQLNLSLDQEIIFNVVLKTVTEQDSSGQAIKIYNNDLTACMWAVDNNILPYTPAA NSMETLGPYPWKPTIASPYRYFCVDRDLSVTYENQGTIEHNTWGTGPKGNSQPFIT ENTQIITLRTGDFEATCTGYFDTNPVKLTHWTQTNROLGQPPILLSTPEADTDAGTL TAQSGRHGATQMEVNVWSEAIRTRPAQVGFQCPHNDPEASRAGPAPKVPADVTQGV DREANGSVYISYKQHGEMWAHPAPERIYTWETNFGSGRDRDGFIOQAPLVVPPP LNGILTANAPIGTKNDIHFNSVNSYGLPTAFPSHPSPVYVYQGIQWKELDLEHKPLRLH ITAPFVCKNAPFGQMLVRLGNLTDQYDPNGATLSRIYTYGTFWKGKLTWRKURAN TTWNPVYQVSDVDNGNSYMSVTKWLPTATGNMQSVPLITRVARNTY" 2287. .2355 /gene="VP1" /codon_start=1 /product="unknown protein" /protein_id="AAA69568.1" /db_xref="GI:825478" /translation="MAPPKRAKRGKGLRDGLWLVGY" 2287. .2317 /gene="VP1" /number=2 2318. .2399 /gene="VP1" /note="alternative intron" 2333. .2362 /gene="VP1" /note="ORF3; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69570.1" /db_xref="GI:825479" /translation="MVGWNGINV" 2355. .2399 /gene="VP1" /note="ORF2; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69571.1" /db_xref="GI:825480" /translation="MFNYLFVRPEITWF" 2400. .>4558 /gene="VP1" /number=2 2795. .4558 /gene="VP2" 2795. .4558 /gene="VP2" /codon_start=1 /protein_id="AAA69572.1" /db_xref="GI:332292" /translation="MSDGTSPQDGGNAVHSAARVERAADPGGGGGGGGGVGVST GSYDQTHRYFLGDGWEVETALATRLVHLNMPKSENYCRIRVHNTTDSYKGNMAKDD AHEQIWTWPSLVADANAWVLQPSDWQYICNTMSQLNLSLDQEIIFNVVLKTVTEQDS GQOAIKIYNNDLTACMWAVDNNILPYTPAANSMTETLGFYFKPTIAGPYRYFCVD RLSVTYENQGTIEHNTWGTGPKGNSQPFITENTQIITLRTGDFEATCTGYFDTNP VKLITWQTNROLGQPPILLSTPEADTDAGTLTAQSGRHGATQMEVNVWSEAIRTRPA QVGFQCPHNDPEASRAGPAPKVPADVTQGVDRANGSVYISYKQHGEMWAHPAPERI PERYTWETNFGSGRDRDGFIOQAPLVVPPPILNLTANPIGTKNDIHFNSVNSYGLPTAF PSHPSPVYVYQGIQWKELDLEHKPLRLHITAPFVCKNAPFGQMLVRLGNLTDQ YDPNGATLSRIYTYGTFWKGKLTWRKURANTTWNPVYQVSDVDNGNSYMSVTKWLPTATGNMQSVPLITRVARNTY"

ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948; Conservative	0; Mismatches 71; Indels 0; Gaps 0;	
Qy	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60	1021 TTTGACCTTAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db	262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAAACCAACCACTGGTTTAAAGGAAAA 321	1282 TTTGACCTTATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1341
Qy	61 AGTAACCAAGGAAGTGTCTCATTTGTTTAAAAAATGAAATGTTCAACTGGAATGGAATA 120	1081 ACAAGAGCCTGCAGAAATTTTGTCTTTTCATGGCTGGAACATATCTTAAAGTTTGCATGCT 1140
Db	322 AGTAACCAAGGAAGTGTCTCATTTGTTTAAAAAATGAGATGTTCAACTTAATGGAATA 381	1342 ACAGAACTCTGCAAGATTTTGTCTTTTCATGGCTGGACATATGTTAAAGTTTGCATGCT 1401
Qy	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGAAGAGCTGGAATCTTTACAA 180	1141 ATTTGCTGTGTTTAAACACAGAGAGGCAAAAGAAATACTGTTTATTTTCATGAGACCA 1200
Db	382 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGAGCAGCTGGAATCTTTACAA 441	1402 ATTTGCTGTGTTCTAAACACAGAGAGGCAAAAGAAATACTGTTTATTTTCACGAGCA 1461
Qy	181 CGAGGAGCGAATCTTGGGACCAAGCGAGGACATGGAATGGGNAACCAAGTGGAT 240	1201 GCAGACACAGGCAAAATCTATTTATTCACAAGCCATAGCAACAGCAGTTGSCATTTGGT 1260
Db	442 CGAGGAGCGAATCTTGGGACCAAGCGAGGACATGGAATGGGNAATCTACAGTGGAT 501	1462 GCAGTACAGGCAAAATCTATTTATTCACAAGCCATAGCAACAGCAGTTGTTAAATTTGGT 1521
Qy	241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTTGA 300	1261 TGTATAATGCAAGCAATGTAAACTTTTCCATTTTAAATGCTGTACCAACAGAACTTTGATT 1320
Db	502 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTTGA 561	1522 TGTATAATGCAAGCTAATGTGAATTTTCCATTTTAAATGCTGTACCAACAGAACTTTGATT 1581
Qy	301 GTGCTTTAACCAAAAGAAATATTTCTCTGTGATGTTAAATTTGTTTGTGCAACATGAATG 360	1321 TGGGTAGAAAGAGCTGTTAACTTTTGACACAGCAAGTAAACCCAGTTTAAAGCCATTTGCTCT 1380
Db	562 GTGCTTTAGCAAAAAATATAGCTCTCTGTGATGTTTACTTTGTTGTCAGCATGAATG 621	1582 TGGGTAGAAAGAGCTGTTAACTTTTGACACAGCAAGTAAACCCAGTTTAAAGCCATTTGCTCT 1641
Qy	361 GGAAGAGCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT 420	1381 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGGAGCAAAAAGATTTGAACCAACCA 1440
Db	622 GGAAGAGCAAGCTGGCACTGCCATGTACTAATTTGGAGGCAAGAGCTTTAGTCAAGCT 681	1642 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGGAGCAAAAAGATTTGAACCAACCA 1701
Qy	421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480	1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGTAAACCCAGTTTGAACCAACCA 1500
Db	682 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 741	1702 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGTAAACCCAGTTTGAACCAACCA 1761
Qy	481 TGTAAATGTCACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540	1501 GAAACACATCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTGCCT 1560
Db	742 TGTAAATGTCACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 801	1762 GAAACACATCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTGCCT 1821
Qy	541 GAGTGGTTTACTTACTTATAGCATATAGCAAAACCAAAAGAGCTATACCAAGTGT 600	1561 GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTTGGTTGGTTGTTAAAG 1620
Db	802 GAGTGGTTTACTTACTTATAGCATATAGCAAAACCAAAAGAGCTATACCAAGTGT 861	1822 GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTTGGTTGGTTGTTAAAG 1881
Qy	601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 660	1621 AATGGTTTCACTTACCATGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGTGATTGG 1680
Db	862 GTTCTTTTGGAAATGATGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 921	1882 AATGGTTTCACTTACCATGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGTGATTGG 1941
Qy	661 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 720	1681 TCAGAAAACTGGCGGAGCAAAAGTGCACAACTCTTATAAATTTACTAGGTTTCGGCACGC 1740
Db	922 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 981	1942 TCAGAAAACTGGCGGAGCAAAAGTGCACAACTCTTATAAATTTACTAGGTTTCGGCACGC 2001
Qy	721 AAGAGAGCGAGCGCATCTAGTGAGCAACTATATACATGATGACATGCGGCCAGAAACG 780	1741 TCACCAATTCAGCACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800
Db	982 AAGAGAGCGAGCGCATCTAGTGAGCAACTATATACATGATGACATGCGGCCAGAAACG 1041	2002 TCACCAATTCAGCACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 2061
Qy	781 GTTGAACCACTAGTACCACTGCGAGGAACCTAAGCGGCGAGATTCAACTTAATAA 840	1801 GCATCGGATCTCGAGACCTGCTTTAGAGCTTTAGAGCTTTGAGACACCAAAATCTCTCTGTTGG 1860
Db	1042 GTTGAACCACTAGTACCACTGCGAGGAACCTAAGCGGCGAGATTCAACTTAATAA 1101	2062 GCATCGGATCTCGAGACCTGCTTTAGAGCTTTGAGAGCTTTGAGACACCAAAATCTCTCTGTTGG 2121
Qy	841 GAAATTTCTTATTAATACTACATTTAAAGAGCTGGTGATATAAAGAGTAACTCACCAGAG 900	1861 GGCATCTGAGAAACCCAGAACTCTGGGGAAGCTGGTTCCAAAGCTTCCAAAGATGTTCAA 1920
Db	1102 GAAATTTCTTATTAATACTACATTTAAAGAGCTAGTGATATAAAGAGTAACTCACCAGAA 1161	2122 GGCATCTGAGAAACCCAGAACTCTGGGGAAGCTGGTTCCAAAGCTTCCAAAGATGTTCAA 2181
Qy	901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGCTCAACCGAGTGGAGAA 960	1921 CTGAGCCCAACTTTGGTTGAGAGATTCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAAACCG 1980
Db	1162 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGCTCAACCGAGTGGAGAA 1221	2182 CTGAGCCCAACTTTGGTTGAGAGATTCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAAACCG 2241
Qy	961 AACCTGCTGAAAAATAGCTAGAGATTTGTACACTAATCTAGCCAGCAACCAACAGCA 1020	1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db	1222 AACCTGCTGAAAAATAGCTAGAGATTTGTACACTAATCTAGCCAGCAACCAACAGCA 1281	2242 TTGAAGAGAGACTTCAGCGAGCGCTGAACTTGGACTAA 2280

RESULT 9

MOU34253

LOCUS

4764 bp

DNA

linear

VRL 21-AUG-1996

DEFINITION	Mouse parvovirus 1b DNA.
ACCESSION	U34253
VERSION	U34253.1 GI:1464793
KEYWORDS	
SOURCE	Mouse parvovirus 1b
ORGANISM	Mouse parvovirus 1b
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
AUTHORS	1 (bases 1 to 4764) Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.
TITLE	Molecular characterization of newly recognized rodent parvoviruses
JOURNAL	J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE	96201434
PUBMED	8609486
REFERENCE	2 (bases 1 to 4764) Besselsen, D.G.
AUTHORS	Direct Submission
JOURNAL	Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES	Location/Qualifiers
Source	1..4764 /organism="Mouse parvovirus 1b" /mol_type="genomic DNA" /specific_host="Mus musculus" /db_xref="taxon:42841"
ORIGIN	
Query Match	94.2%; Score 1902.2; DB 14; Length 4764;
Best Local Similarity	96.4%; Pred. No. 0;
Matches 1946; Conservative	0; Mismatches 73; Indels 0; Gaps 0;
QY	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
DB	122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 181
QY	61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
DB	182 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAACTGAGATGTTCAACTGAATGAAAA 241
QY	121 GATATCGGATGGATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 180
DB	242 GATATCGGATGGAAATTAATACAAAAAGGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 301
QY	181 CGAGGAGCGGAAACTACTTTGGGACCAAGCAGGACATGGAATGGGAAACCACAGTGGAT 240
DB	302 CGAGGAGCGGAAACTACTTTGGGACCAAGCAGGACATGGAATGGGAACTACAGTGGAT 361
QY	241 GAAATGACCAAAAGCAAGTATTCATTTTGTATTTTGTGTTTAAAAAATGTTTATTTGAA 300
DB	362 GAAATGACCAAAAGCAAGTATTCATTTATGACTCTTTAGTTAAAAAATGTTTGTGAA 421
QY	301 GTGCTTAAACAAAGATATATTTCTGCTGATGTTAATTTGTTTGTGCAATGATGG 360
DB	422 GTGCTTAGCAAAAGATATAGTCTCTGCTGATGTTTACTTTGTTGTGAGCATGAATGG 481
QY	361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGACATTTAGTCAAGCT 420
DB	482 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGACAGCACTTTAGTCAAGCT 541
QY	421 CAAGGAAATGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAGCAGCC 480
DB	542 CAAGGAAATGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGTTAAGCAGCC 601
QY	481 TGTATGTGCACTACACAGCTGAAGAAATTAACCTAAGAAATAGCAGAGCAAT 540
DB	602 TGTATGTGCACTACACAGCTGAAGAAATTAACCTAAGAAATAGCAGAGCACT 661
QY	541 GAGTGGTTTACTTCTACTTACTTATAAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGT 600
DB	662 GAGTGGTTTACTTCTACTTACTTATAAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGT 721
QY	601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAAAGAAATAAGCACTAGT 660
DB	722 GTTCTTTTGGAAACATGATGCTTACTATTTTAAACCAAGAAATAAGCACTAGT 781
QY	661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTACTTTTAA 720
DB	782 CGCCCAAGGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTACTTTTAA 841
QY	721 AAAGAGGCGAGCGCCCATCTAGTGAAGCAAACTATACACTGATGACATCGGCGCAGAAACG 780
DB	842 AAAGAGGCGAGCGCCCATCTAGTGAAGCAAACTATACACTGATGACATCGGCGCAGAAACG 901
QY	781 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGGAGAAATCAAACTAAAAA 840
DB	902 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGGAGAAATCAAACTAAAAA 961
QY	841 GAAGTTTCTATTAAACCTACACTTAAAGAGCTGGTGCATAAAGAGTAGTACCTCACCAGAG 900
DB	962 GAGGTTTCTATTAAACCACTTAAAGAGCTGGTGCATAAAGAGTAGTACCTCACCAGAA 1021
QY	901 GACTGATGATGATGAGCAGCAGACAGTTCATTTGAAATGATGGCTCAACAGAGTGGAGAA 960
DB	1022 GACTGATGATGATGAGCAGCAGACAGTTCATTTGAAATGATGGCTCAACAGAGTGGAGAA 1081
QY	961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTCTAGCCAGAAACCAACAGCA 1020
DB	1082 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTCTAGCTAGAAACCAACAGCA 1141
QY	1021 TTTTGACTTAAATTTAGAAAAAGCTGAAACCAACCACTAAACCACTTTTCACTGCCTGAC 1080
DB	1142 TTTTGACTTAAATTTAGAAAAAGCTGAAACCAACCACTAAACCACTTTTCACTGCCTGAC 1201
QY	1081 ACAAGAGCTGCAGAAATTTTTCATGCTGGAACTATGTTTAAAGTTTGCCATGCT 1140
DB	1202 ACAAGAGCTGCAGAAATTTTTCATGCTGGAACTATGTTTAAAGTTTGCCATGCT 1261
QY	1141 ATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTATGAGACCA 1200
DB	1262 ATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTTATGAGACCA 1321
QY	1201 GCAGCAGACAGGCAAACTTATTATTCACAAGCCATAGCACAAGAGCTTGGCAATGTTGT 1260
DB	1322 GCAGTACAGGCAAACTTATTATTCACAAGCCATAGCACAAGAGCTTGGTAAATGTTGT 1381
QY	1261 TGCTATAATGCAAGCAATGTAACCTTTCCATTTAATGACTGTACCAACAGAACTTGAT 1320
DB	1382 TGCTATAATGCAAGCAATGTAACCTTTCCATTTAATGACTGTACCAACAGAACTTGAT 1441
QY	1321 TGGGTAGAAGAGCTGGTAACTTTTGAACAGCAAGTAAACAGCTTTTAAAGCCATTTGCTCT 1380
DB	1442 TGGGTAGAAGAGCTGGTAACTTTTGAACAGCAAGTAAACAGCTTTTAAAGCCATTTGCTCT 1501
QY	1381 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGCAAAACAGATTGCAACCAACCA 1440
DB	1502 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGCAAAACAGATTGCAACCAACCA 1561
QY	1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA 1500
DB	1562 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA 1621
QY	1501 GAAACACTCAACCAATCAGAGACAGATGCTTAACTTCACTTCAACATACCTTTGCTCT 1560
DB	1622 GAAACACTCAACCAATCAGAGACAGATGCTTAACTTCACTTCAACATACCTTTGCTCT 1681
QY	1561 GGTGACTTTGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
DB	1682 GGTGACTTTGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1741
QY	1621 AATGTTACCAATCTACCAATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGATGG 1680
DB	1742 AATGTTACCAATCTACCAATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGATGG 1801
QY	1681 TCAGAAAACTGGGCGGAGCAAGGTGCAACTCTCTATTAATTTACTAGTTTCGGCAGCC 1740

```
Db 1802 ACAGAAAACCTGGCGGACGAGGTCCGACTCCTATATAAATCACTAGTTCGGCACGC 1861
Qy 1741 TCACCAATTCACGACACCGAAAAGTAGCCCTCTCAGCAGAACATATGCACTAACCTCACTT 1800
Db 1862 TCACCAATTCACGACACCGAAAAGTAGCCCTCTCAGCAGAACATATGCACTAACCTCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAATACTCCTGTTGCG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAATACTCCTGTTGCG 1981
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAAGATGGTCAA 1920
Db 1982 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAAGATGGTCAA 2041
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGCAGAACCG 1980
Db 2042 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGCAGAACCG 2101
Qy 1981 TTGAAGNAGACTTCAGGAGCGCTGAACCTGACCTAA 2019
Db 2102 TTGAAGNAGACTTCAGGAGCGCTGAACCTGACCTAA 2140

RESULT 10.
PAMVMI
LOCUS
DEFINITION Mouse parvovirus minute virus immunosuppressive variant genome (=
PAMVMI).
ACCESSION X02481
VERSION X02481.1 GI:60918
KEYWORDS coat protein; genome; origin of replication; overlapping genes;
terminal repeat; unidentified reading frame.
SOURCE Mouse minute virus
ORGANISM Mouse minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.
TITLE DNA sequence comparison between two tissue-specific variants of the
autonomous parvovirus, minute virus of mice
Nucleic Acids Res. 13 (10), 3617-3633 (1985)
JOURNAL 85242059
MEDLINE 3855242
PUBMED
COMMENT For the fibroblast-specific strain (MVMp) sequence see <PAMW2>.
The genomes of MVMp and MVMi (immunosuppressive variant) have more
than 96% of their sequence in common.
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES
source 1..5087
/organism="Mouse minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
misc_feature 1..116
/notes="terminal hairpin region"
promoter 176..183
/notes="TATA box"
gene 2405..4558
/genes="vpi"
CDS <2405..4558
/genes="vpi"
/function="coat protein"
/codon_start=1
/product="vpi protein"
/protein_id="CAB46507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="VPPGYKLGPNLSLDQGEPTNPSDAAKHEHDAYDQYIKSGKNP
YLYPSAADORFIQDQKADGWGKGVGHYFFRTKRAFAPKLATDSEPTGSGVSAGKRT
RPPAYIFINQARAKKLTSSAAQSSOTMSDGTSSQPDGNGVHSAARVERAADPGGS
GGGSGGGGVGVSTQDTHYRFLDGHVEITATALRLVHLNMPKSENYCRIRVHN
TTDTSVKNMAKODAHQEIWTPSLVDANANGWLQPSDQWYICNTMSQLNLSLQDE
IFNVLVKTVEQDSGGQAIKIYNNDLTACMVAVDSNNILPYPANSMETLGFYPMK
PTIASPYRYFCVDRDLSTVYENQEGTIEHNVMGTPKGMNSQFFTIENQOITLLRTG
```

```
DEFAFGTYVVDNTPVKLTHTWOTNRQLGQPLLSTFPPEADTDAGTLTAQSGSRHGATQM
EYNNVSEAIRTEPAQVGCOPHNDPEASAGPAAPKVPADVTQGVDRANGSVRSY
GKHQENWAAHGFAFERITWDETNGSGRDRDGFQISAPLVVPPPLNGILTNAPIG
TKMLDHFNSVFNSTGYDPAFSPHSVYIQQQIWKDELDELHKPRLHITAPFVCKNNAP
GQMLVRLGNLTQDTPNGATLSRIVTGYTFWKGKLTMRKLRANTTNPNVYQVNSVE
DNGNSYMSVTWMLPTATGMQSVPLITRPARNTY"
2792..4558
/genes="vpi2"
CDS <2792..4558
/function="coat protein"
/codon_start=1
/product="vpi2 protein"
/protein_id="CAB46508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="TMSDGTSGDNGVHSAARVERAADPGGSGGGSGGGGVGS
TGSYDNOHYRFLDGHVETATALRLVHLNMPKSENYCRIRVHNITDTSVKNMAK
DAHEQIWTWPSLVLDANANGWLQPSDQWYICNTMSQLNLSLQDEIFNVLVKTVEQD
SGGQAIKIYNNDLTACMVAVDSNNILPYPANSMETLGFYPMKPTIASPYRYFCV
DRDLSTVYENQEGTIEHNVMGTPKGMNSQFFTIENQOITLLRTGDEFAFGTYVVDN
PVKLTHTWOTNRQLGQPLLSTFPPEADTDAGTLTAQSGSRHGATQMEVNNVSEAIRTEP
AQVGCOPHNDPEASAGPAAPKVPADVTQGVDRANGSVRSYQKHQENWAAHGFA
AERYTWDHETNGSGRDRDGFQISAPLVVPPPLNGILTNAPIGTKMLDHFNSVFN
YGLPTAFSPHSVYIQQQIWKDELDELHKPRLHITAPFVCKNNAPQMLVRLGNLTQD
QYDNGATLSRIVTGYTFWKGKLTMRKLRANTTNPNVYQVNSVEDNGNSYMSVTWML
PTATGMQSVPLITRPARNTY"
4603..4608
/misc_feature
/notes="polyadenylation signal"
4821..4826
/misc_feature
/notes="polyadenylation signal"
4878..5087
/misc_feature
/notes="terminal hairpin region"

ORIGIN
Query Match 94.2%; Score 1902.2; DB 14; Length 5087;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTCTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTCTTGGGAGCAACCAACTGGTTAAAGGAAAA 321
Qy 61 AGTAACCAAGAAAGTGTCTCATTTGTTTAAATAATGAAATTTCAACTGAATGGAATA 120
Db 322 AGTAACCAAGAAAGTGTCTCATTTGTTTAAATAATGAAATTTCAACTGAATGGAATA 381
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGGAATCTTTACAA 180
Db 382 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGGAATCTTTACAA 441
Qy 181 CGAGGAGCGGAAACTACTTGGGACCAAGAGAGCTGGAATGGAATGGAATGGAATGGAAT 240
Db 442 CGAGGAGCGGAAACTACTTGGGACCAAGAGAGCTGGAATGGAATGGAATGGAATGGAAT 501
Qy 241 GAAATGACCAAAAAGCAAGTATTTCATTTTGGTATTTTGGTATTTTAAATAATGTTTATTTGAA 300
Db 502 GAAATGACCAAAAAGCAAGTATTTCATTTTGGTATTTTGGTATTTTAAATAATGTTTATTTGAA 561
Qy 301 GTGCTTAAACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTCGCAACATGATGAG 360
Db 562 GTGCTTAAACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTCGCAACATGATGAG 621
Qy 361 GGAAGAGACCAAGCGTGGCACTGCCATGTAATTTGGAGGAAGAGCTTTTAGTCAAGCT 420
Db 622 GGAAGAGACCAAGCGTGGCACTGCCATGTAATTTGGAGGAAGAGCTTTTAGTCAAGCT 681
Qy 421 CAAGGGAATAATGTTGGAGAGAGGCACTAAATTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480
Db 682 CAAGGGAATAATGTTGGAGAGAGGCACTAAATTTTACTTGGAGCAGATGTTTGGTAAACAGCC 741
Qy 481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAAGAGAAATAGCAGAGAACAT 540
```







OTKKEYSIKITLXELVHKVTSPEDEMMQPDYSIEMMAQPGGNNILKNLTBICITLTL		ARTKFAFLLEKAEKTSKUNFSLPDRICKIFAPFGWNVIKVKCHACCVLNKQGGKR	
NTVLFHPGASTGSIIIAQAQAVGNVCYNAANVPFNDCTNKNLIWVEAGNFGQ		QVNFPAICSGQITIDOKGSKQIEPTVIMKTNGNITVVRKIGSEERPEHTQPIRD	
RMLNIHLTHLPDGFGLVDKNWPMICAWLVKNQYOSTWASVCAKWKVQPDWENWAE		PKVPTINSLSGARSPTTPKSPPLSNVALTPLASDLIEDLALPEWSTPNTVPVGTAE	
TQNTGSGSKACODGQLSPWSEIEDLRACFGAEPLKXDFSEPLND"		join(2290..2320,2403..4561)	
gene	/gene="VP1"	/genes="VP1"	
CDS	join(2290..2320,2403..4561)	/gene="vpi"	
/codon_start=1		/product="capsid protein"	
/protein_id="AAA61406.1"		/db_xref="GI:525327"	
/translation="MAPPAKRAKRGWVPPQYKYLPGNSLDQGEPTNPDSAAKEHDE		AYDKYIKSGNPNLYFSADORFIDOTKADKGGVGHVFPRTKFAFAPRLASSEP	
GTSVGIAGKRTKPPAHIFINOARAKKRSLAAQOORTOTMSDGAQPPDGSVAQSSA		RVERAADGPGSGSGSGGVGVSTGSDYNQTHYRFLSDGWVEITAYSTRVHLNMP	
KSENYCRVRHNTNDRTAGNMAKDHAHQIWTPLSLIDSNWGVWFQSDWQFICNN		MSHVNLSLDQELFNVIIVITVTEQNTGAEAVKYIYNNDLTASMMVALDSNNILPYTPAT	
DNQETLGFYKFWKTPSPYRYFNCDSLSVTYDQTSIGVTDMANASGLSOFFTIE		NTORIQLLRGTDFATGYTFEPIKLSHTWQSNRQLGPPPOITDLPADNENATLV	
TRGDRSITQIISGNDVTEATRVRPAQVGFQCPDHNFETSRAGPKVPVVPVANVTQGN		EHDANGSLRYTDKQHQDQWGSNNKERTWDALSYDSGRWADRCFINATPFTSPPAL	
NNILTNSDPIGNKTAIHYQNVSNSYGLTAFPHAPPIYPQGOIWDKELDEHKPRIHA		QAPFVCKNNAPOGLLVRLAPNLTDQDPSNLSLRYVTYTFTFKGLKLKAKLRPNA	
TWNPVYQVSAQYQNEVMSIHKWLPATATGNMQSIPLLSRPVARNTY"			
ORIGIN			
Query Match 94.1%; Score 1899; DB 14; Length 5144;		Best Local Similarity 96.3%; Pred. No. 0;	
Matches 1944; Conservative 0; Mismatches 75; Indels 0; Gaps 0;			
Qy	1	ATGGCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA	60
Db	265	ATGGCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA	324
Qy	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAATA	120
Db	325	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAATA	384
Qy	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTCAGAGGACGAGCTGAAATCTTTACAA	180
Db	385	GATATCGGATGGAATAGTTTACAAAAAGAGCTCAGAGGACGAGCTGAAATCTTTACAA	444
Qy	181	CGAGGAGCGGAACTACTTCTGGACCAAGAGCGAGCATGGAATGGGAAACCAAGTGGAT	240
Db	445	CGAGGAGCGGAACTACTTCTGGACCAAGAGCGAGCATGGAATGGGAACTACAGTGGAT	504
Qy	241	GAAATGACCAAAAAGCAAGTATTCATTTTGTATCTTTGTTTAAAAAATGTTTATTTGAA	300
Db	505	GAAATGACCAAAAAGCAAGTATTCATTTGATGACTCTTTAGTTAAAAAATGTTTATTTGAA	564
Qy	301	GTGCTTAAACAAAGAAATATTTCTCTGGTATGTTAAATTTGGTTTGTGCAACATGAATGG	360
Db	565	GTGCTTAAACAAAGAAATATTTCTCTGGTATGTTAAATTTGGTTTGTGCAACATGAATGG	624
Qy	361	GGAAAGACCAAGGCTGGCACTGCCATGACTAATTTGGAGGAAGGACTTTAGTCAAGCT	420
Db	625	GGAAAGACCAAGGCTGGCACTGCCATGACTAATTTGGAGGAAGGACTTTAGTCAAGCT	684
Qy	421	CAAGGAAATTTGGAGGAAGGCACTAAATTTGTTTACTTGGAGCAGATGGTTGTTAAACAGCC	480
Db	685	CAAGGAAATTTGGAGGAAGGCACTAAATTTTACTTGGAGCAGATGGTTGTTAAACAGCC	744
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAAT	540
Db	745	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAAT	804
Qy	541	GAGTGGGTTACTTACTTACTTATAGCATAGCAACCAACCAAGACTATACCAAGTGT	600
Db	805	GAGTGGGTTACTTACTTACTTATAGCATAGCAACCAACCAAGACTATACCAAGTGT	864

Qy	601	GTTCTTTTGGAAAAATAGATGCTTACTTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
Db	865	GTTCTTTTGGAAAAATAGATGCTTACTTATTTTAACTAAAAAGAAAAATAAGCACTAGT	924
Qy	661	CCACCAAGAGACGGAGGCTATTTCTTACAGTGACTCTGGCTGGAAAACTTAACCTTTTA	720
Db	925	CGCCCAAGAGACGGAGGCTATTTCTTACAGTGACTCTGGCTGGAAAACTTAACCTTTTA	984
Qy	721	AAAGAAAGGAGCGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Db	985	AAAGAAAGGAGCGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	1044
Qy	781	GTTGAAAAACACAGTAAACCACTGCGGAGGAAAACTAAGCGGCGGAGAAATTCAACTAAAAAA	840
Db	1045	GTTGAAAAACACAGTAAACCACTGCGGAGGAAAACTAAGCGGCGGAGAAATTCAACTAAAAAA	1104
Qy	841	GAAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG	900
Db	1105	GAGTTTCTATTAAAAACCACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAA	1164
Qy	901	GACTGATGATGATGAGCGAGACAGTACATTGAAATGATGCTCAACACGAGTGGAGAA	960
Db	1165	GACTGATGATGATGAGCGAGACAGTACATTGAAATGATGCTCAACACGAGTGGAGAA	1224
Qy	961	AACCTGCTGAAAAAATACGCTAGAGATTGTACACTTACCTGACGAGAACCAAAACAGCA	1020
Db	1225	AACCTGCTGAAAAAATACGCTAGAGATTGTACACTTACCTGACGAGAACCAAAACAGCA	1284
Qy	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCGAGCAAACTAACCACTTTTCACTGCCTGAC	1080
Db	1285	TTTGACTTAAATTTAGAAAAAGCTGAAACCGAGCAAACTAACCACTTTTCACTGCCTGAC	1344
Qy	1081	ACAAGAGCTGCGAGAAATTTTGTCTTTCATGGCTGGAACTATCTTAAAGTTTCCCATGCT	1140
Db	1345	ACAAGAGCTGCGAGAAATTTTGTCTTTCATGGCTGGAACTATCTTAAAGTTTCCCATGCT	1404
Qy	1141	ATTGCTGTGTTTTTAAACAGAGAGGAGGAAAAAATACTGTTTTTATTTTCATGGACCA	1200
Db	1405	ATTGCTGTGTTTTTAAACAGAGAGGAGGAAAAAATACTGTTTTTATTTTCATGGACCA	1464
Qy	1201	GCAGCAGAGGCAAACTATTTATTCAGCAAGCCATAGCAGCAAGCTTGGCAATGTTGGT	1260
Db	1465	GCAGCAGAGGCAAACTATTTATTCAGCAAGCCATAGCAGCAAGCTTGGCAATGTTGGT	1524
Qy	1261	TGCTATAATGCGAGCAATGTAACTTTTCAATTTAATGACTGTACCAACCAAGAACTTGATT	1320
Db	1525	TGCTATAATGCGAGCAATGTAACTTTTCAATTTAATGACTGTACCAACCAAGAACTTGATT	1584
Qy	1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAAACAGTTTAAAGCCATTGCTCT	1380
Db	1585	TGGGTAGAAAGAGCTGGTAACTTTTGGACAAACAGTAAAACAGTTTAAAGCCATTGCTCT	1644
Qy	1381	GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGACCAACCA	1440
Db	1645	GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGACCAACCA	1704
Qy	1441	GTCAATCATGACCAAAATGAGAACTTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA	1500
Db	1705	GTCAATCATGACCAAAATGAGAACTTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA	1764
Qy	1501	GAAACACATCAACCAATCAGAGACAGAAATGCTTTAACTTATCATCTAAACATCTTGCCT	1560
Db	1765	GAAACACATCAACCAATCAGAGACAGAAATGCTTTAACTTATCATCTAAACATCTTGCCT	1824
Qy	1561	GGTGACTTTGGTTGGTTGACAAAAATGAAATGGCCCCATGATTTGCTGGTTGGTTAAG	1620
Db	1825	GGTGACTTTGGTTGGTTGACAAAAATGAAATGGCCCCATGATTTGCTGGTTGGTTAAG	1884
Qy	1621	AATGGTTTACCAATCTACCACTGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATTG	1680
Db	1885	AATGGTTTACCAATCTACCACTGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATTG	1944

```
QY 1681 TCAGAAACCTGGCGGAGCCAAAGGTGCCAACCTCTATATAATTTACTAGTTCGGCAGC 1740
Db 1945 ACGGAAACCTGGCGGAGCCGAGGTGCCGACTCTATAAATTCACCTAGTTCGGCAGC 2004
QY 1741 TCACCAATTCAGGACACCGAAAGTACCGCTCTAGCCAGAACATATGCACTAACTCCACTT 1800
Db 2005 TCACCAATTCAGGACACCGAAAGTACCGCTCTAGCCAGAACATATGCACTAACTCCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCAGACACCAATATCTCTCTGTGCG 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCAGACCAATATCTCTCTGTGCG 2124
QY 1861 GGCACTCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 2125 GGCACTCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2184
QY 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTGAGACCGTCTCGTGGCGGAACCG 1980
Db 2185 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTGAGACCGTCTCGTGGCGGAACCG 2244
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2245 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2283

RESULT 12
MOU34254
LOCUS Mouse parvovirus 1c DNA. 4764 bp DNA linear VRL 21-AUG-1996
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS .
SOURCE Mouse parvovirus 1c
ORGANISM Mouse parvovirus 1c
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L.,
Franklin, C.E., Hook, R.R. Jr. and Riley, L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen, D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
Location/Qualifiers
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATCGTCTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60
Db 122 ATGGCTGGAAATCGTCTACTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 181
QY 61 AGTAACCCAGGAAGTTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGGAATGAAAAA 120
Db 182 AGTAACCCAGGAAGTTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGGAATGAAAAA 241
QY 121 GATATCGGATGGAATGTTACAAAAAGAGCTCGAGGAGCAGCTGAAATCTTTTACAA 180
Db 242 GATATGGAATGGAATGTTACAAAAAGAGCTCGAGGAGCAGCTGAAATCTTTTACAA 301
QY 181 CGAGGAGCGGAAACTACTTGGGACCAAGCGGAGCATGGAATGGGAAACCACTGGAT 240
```

```
Db 302 CGAGGAGCGGAAACTACTTGGGACCAAGCGGAGCATGGAATGGGAATCTACAGTGGAT 361
QY 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTTTGGTTTAAATAATGTTTATTTGAA 300
Db 362 GAAGTGACCAAAAGCAAGTATTCATTTATGACTCTTTAGTTTAAATAATGTTTATTTGAA 421
QY 301 GTGCTTAAACAAAGAAATATATTTCTCGTGTATGTTAAATGGTTTGTCAACATGAATGG 360
Db 422 GTGCTTAAACAAAGAAATATCTCTCGTGTATGTTAAATGGTTTGTGAGCATGAATGG 481
QY 361 GGAAGAAGCAAGGCTGGCACTGCCATGTAATAATTGGAGGAAAGACTTTTGTCAAGCT 420
Db 482 GGAAGAAGCAAGGCTGGCACTGCCATGTAATAATTGGAGGCAAGACTTTTGTCAAGCT 541
QY 421 CRAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGTTTAAACAGCC 480
Db 542 CRAAGGAAATGTTGGAGAGGCAAGTAAAGTGTCTGAGCAGATGTTGTTTAAACAGCC 601
QY 481 TGTAAATGTGCAACTAACCCAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 540
Db 602 TGTAAATGTACAGCTAACCCAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 661
QY 541 GAGTGGTTTACTCTACTTACTTAAAGCAATAGCAAAACCAAAAGACTATATACCAAGTGT 600
Db 662 GAATGGTTTACTCTTGTCTACTTAAAGCAATAGCAAAACCAAAAGACTATATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATTCCTTACTATTTTAACTTAAAGAAATAAGCACTAGT 660
Db 722 GTTCTTTTGGAAACATGATTCCTTACTATTTTAACTTAAAGAAATAAGCACTAGT 781
QY 661 CCACCAAGAGAGCGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db 782 CCGCCAAAGGAGCGAGGCTTATTTCTTAGTACTCTGGCTGGAAAACTAACTTTTAA 841
QY 721 AAAGAAGGAGCGGCGCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGGAGAAACG 780
Db 842 AAAGAAGGCGGAGCGGCGCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGGAGAAACG 901
QY 781 GTTGAACACACAGTAACTACCTCGCGAGGAAACTAAGCGCGGAGAACTTCAAACTTAAAAA 840
Db 902 GTTGAACACACAGTAACTACCTCGCGAGGAACTAAGCGCGGAGAACTTCAAACTTAAAAA 961
QY 841 GAAGTTTCTTATTTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCCACAGAG 900
Db 962 GAGGTTTCTTATTTAAAAACCACTTAAAGAGCTAGTGCAATAAGAGAGTAACTCCACAGAA 1021
QY 901 GACTGGATGATGTCAGCGCAGACAGTTACATTGAAATGATGCTCAACAGCTGGAGAA 960
Db 1022 GACTGGATGATGTCAGCGCAGACAGTTATATTTGAAATGATGCTCAACAGCTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAAAATACGCTAGAGATCTGTACACTTAACTCTAGCTAGAGAACCAAAACAGCA 1141
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAACCAAACTAAACCACTTTTCACTGCTGAC 1080
Db 1142 TTTGACTTAAATTTAGAAAAAGCTGAAACCAACCAAACTAAACCACTTTTCACTGCTGAC 1201
QY 1081 ACAAGGCTGCGAGAAATTTTGTCTTTTCTGCTGGAAGTATGTTAAAGTTTGGCATGCT 1140
Db 1202 ACAAGAACCTGCAAGATCTTTGCTTTTCTGCTGGAAGTATGTTAAAGTTTGGCATGCT 1261
QY 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTTCATGGACCA 1200
Db 1262 ATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTTCATGGACCA 1321
QY 1201 GCCAGACAGGCAATCTATTTATTCAGAACGATGACAGCAAGAGCTGGGCAATGTTGCT 1260
Db 1322 GCCAGTACAGGCAATCTATTTATTCAGAACGATGACAGCAAGAGCTGGGCAATGTTGCT 1381
QY 1261 TGCTATATGACGCAATCTTAACTTTTCACTTTATGATGTTTACCAACAGAACTTGATT 1320
```

Db	1382	TGTTATTAATGCAGCAAAATGTGAACCTTTCCAATTCAATGACTGTGTACCAACAGAACTTAATT	1441		
Qy	1321	TGGGTAGAAGAGCTGTGTAACTTTGGACAGCAAGTAAACAGATTTAAAGCAATTTGCTCT	1380		
Db	1442	TGGGTGGAAGAGCTGTGTAACTTTGGACAGCAAGTAAACAGATTTAAAGCAATTTGCTCT	1501		
Qy	1381	GGTCAAACTATTGTCATTTGATTCAAAAAGGAAAAAGGCGAGCAAAACAGATTGAACCAACCA	1440		
Db	1502	GGTCAAACTATTGTCATTTGATTCAAAAAGGAAAAAGGCGAGCAAGTGAACCAACCA	1561		
Qy	1441	GTCAATCATGACCAACAATGAGAACATTTACAGTGGTTCAGAAATAGGCTGCGAGAGAAACCA	1500		
Db	1562	GTCAATCATGACCAACAATGAGAACATTTACAGTGGTTCAGAAATAGGCTGCGAGAGAAACCA	1621		
Qy	1501	GAACACACTCAACCAATFAGAGACAGATGCTTAAACATTCATCTAACACATACATCTTGCCT	1560		
Db	1622	GAACACACTCAACCAATFAGAGACAGATGCTTAAACATTCATCTAACACATACATCTTGCCT	1681		
Qy	1561	GGTGACTTTGGTTGGTTGACAAACATGANTGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620		
Db	1682	GGTGACTTTGGTTGGTTGACAAACATGANTGGCCCATGATTTGTGCTTGGTTGGTAAAG	1741		
Qy	1621	AATGGTTTACCAATCTACCTGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATGG	1680		
Db	1742	AATGGTTTACCAATCTACCTGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATGG	1801		
Qy	1681	TCAGAAACTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGTTTCGGCAGCG	1740		
Db	1802	ACGAAAACTGGGCGGAGCCGAAGTGCACCTCTATAAATTTCACTAGTTTCGGCAGCG	1861		
Qy	1741	TCACCAATTCAGCACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCATTAACCTCACTT	1800		
Db	1862	TCACCAATTCAGCACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCATTAACCTCACTT	1921		
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGACACCAACATCTCTGTTGG	1860		
Db	1922	GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGACACCAACATCTCTGTTGG	1981		
Qy	1861	GGCACTGCGAAGAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGGAACCG	1920		
Db	1982	GGCACTGCGAAGAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGGAAGTGTCAA	2041		
Qy	1921	CTGAGCCCACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGCAGAACCG	1980		
Db	2042	CTGAGCCCACTGGTTCAGAGATCGAGGAGGATTTGAAAGCGTGTCTCGGTGCGGAACCG	2101		
Qy	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019		
Db	2102	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2140		
RESULT 13					
HOU34255					
LOCUS	HOU34255	4773 bp	DNA linear VRL 21-AUG-1996		
DEFINITION	Hamster parvovirus DNA.				
ACCESSION	U34255				
VERSION	U34255.1	GI:1464792			
KEYWORDS					
SOURCE	Hamster parvovirus				
ORGANISM	Hamster parvovirus				
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.				
AUTHORS	1 (bases 1 to 4773) Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.				
TITLE	Molecular characterization of newly recognized rodent parvoviruses				
JOURNAL	J. Gen. Virol. 77 (Pt 5), 899-911 (1996)				
MEDLINE	96201434				
PUBMED	8609486				
REFERENCE	2 (bases 1 to 4773)				
AUTHORS	Besselsen, D.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia W213				

FEATURES		Veterinary Medicine Building, Columbia, MO 65211, USA	
source		Location/Qualifiers	
		1..4773	
		/organism="Hamster parvovirus"	
		/mol_type="genomic DNA"	
		/specific_host="Mesocricetus auratus"	
		/db_xref="taxon:42843"	
ORIGIN		Query Match ^ 92.2%; Score 1862.2; DB 14; Length 4773;	
		Best Local Similarity 95.1%; Pred. No. 0;	
		Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;	
Qy	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Db	122	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCACTGGTTAAAGGAAA	181
Qy	61	AGTAACACAGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGA	120
Db	182	AGTAACACAGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGA	241
Qy	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA	180
Db	242	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGAGCTGAAATCTTTACAA	301
Qy	181	CGAGGAGCGGAAACTACTCTTGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT	240
Db	302	CGAGGAGCGGAAACTACTCTTGGACCAAGCGAGGACATGGAATGGGAAATCTTCACTGGAT	361
Qy	241	GAAATGACCAAAAGCAAGTATTCATTTGTTTGGTTTAAATGTTTAAATGTTTAA	300
Db	362	GAAATGACCAAAAGCAAGTATTCATTTGTTTGGTTTAAATGTTTAAATGTTTAA	421
Qy	301	GTGCTTAAACAAAGAAATATTTCTCTGGTGTATGTTAAATGTTTGTGCAACATGAATGG	360
Db	422	GTGCTGAGTACAAAGAAACATAGCACCTAGTGTGTTTGTGTTGTACAGCATGATGG	481
Qy	361	GGAAAGACCAAGGCTGGCACTGCTAAATTTGGAGGAAAGCACTTTAGTCAAGCT	420
Db	482	GGAAAGACCAAGGCTGGCACTGCTAAATTTGGAGGAAAGCACTTTAGTCAAGCT	541
Qy	421	CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGGTAAACAGCC	480
Db	542	CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGGTAAACAGCC	601
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
Db	602	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAGATAGCGGAGACCA	661
Qy	541	GAGTGGTTTACTCTTACTTATAGCATAAGCAACCAACCAAGACATATACCAAGTGT	600
Db	662	GAGTGGTTTACTCTTACTTATAGCATAAGCAACCAACCAAGACATATACCAAGTGT	721
Qy	601	GTCTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATTAAGCACTAGT	660
Db	722	GTCTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATTAAGCACTAGT	781
Qy	661	CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACTTAACTTTTA	720
Db	782	CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACTTAACTTTTA	841
Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAACATATACACTGATGACATGCGGCGAGAACG	780
Db	842	AAAGAGGCGAGCGCCATCTAGTGAGCAACATATACACTGATGACATGCGGCGAGAACG	901
Qy	781	GTTGAAACACAGTAAACCACTGCGGAGGAACTAAGCGGCGGAGATTCAACTTAA	840
Db	902	GTTGAAACACAGTAAACCACTGCGGAGGAACTAAGCGGCGGAGATTCAACTTAA	961
Qy	841	GAGTGTCTTATTAAACACTACACTTAAAGAGCTGGTGCATATAAGAGAGTAACTCACCAGAG	900
Db	962	GAGTGTCTTATTAAACACTACACTTAAAGAGCTGGTGCATATAAGAGAGTAACTCACCAGAA	1021

```
Qy 901 GACTGGATGATGTCAGCAGACAGCTTACATTGAATGATGCTCAACAGGTGGAGAA 960
Db 1022 GACTGGATGATGTCAGCAGCAGAGTACATTGAATGATGCTCAACAGGTGGAGAA 1081
Qy 961 AACCTGCTGAAATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1141
Qy 1021 TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTTCACCTGCCTGAC 1080
Db 1142 TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTTCACCTGCCGAC 1201
Qy 1081 ACAAGAGCTGCAAGATTTTGGCTTTTCATGGCTGGAATATGTTAAAGTTTCCCATGCT 1140
Db 1202 ACAAGAGCTGCAAGATTTTGGCTTTTCATGGCTGGAATATGTTAAAGTTTCCCATGCT 1261
Qy 1141 ATTGCTGTGTTTAAACAGACAGAGGCGGCAAAAGAAATACTGTTTATTTATGACACCA 1200
Db 1262 ATTGCTGTGTTTAAACAGACAGAGGCGGCAAAAGAAATACTGTTTATTTTCATGACCA 1321
Qy 1201 GCCAGACAGCGCAATCTATTATTGCAACAGCCATAGCAACAGCTTGGCAATGTTGGT 1260
Db 1322 GCCAGACAGCGCAATCTATTATTGCAACAGCCATAGCAACAGCTTGGTAATGTTGGT 1381
Qy 1261 TGCTATAATCCAGCCAAATGTAATCTTTCCATTTAATGACTGTACCAACAGAACTTGAT 1320
Db 1382 TGCTATAATCCAGCAATGTAATCTTTCCATTTAATGACTGTACCAACAGAACTTGAT 1441
Qy 1321 TGGGTAGAAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
Db 1442 TGGGTAGAAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1501
Qy 1381 GGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACACCA 1440
Db 1502 GGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACACCA 1561
Qy 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1621
Qy 1501 GACACACTCAACCAATCAGACAGACAGATGCTTAACTTATCATCTACACATACCTTGCCT 1560
Db 1622 GACACACTCAACCAATCAGACAGACAGATGCTTAACTTATCATCTACACATACCTTGCCT 1681
Qy 1561 GGTGACTTTGGTTGGTTGCAAAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG 1620
Db 1682 GGTGACTTTGGTTGGTTGCAAAACATGATGGCCCATGATTTGCTGGTTGGTAAAG 1741
Qy 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATGG 1680
Db 1742 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATGG 1801
Qy 1681 TCAGAAACTGGGGGAGCCAAAGGTGCCAACTCTTATAATTTACTAGTTTGGCACGC 1740
Db 1802 ACGGAAACTGGGGGAGCCAAAGGTGGCGACTCTCTATAAATTTCACTAGTTTGGCACGC 1861
Qy 1741 TCACCAATTTCAGCACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAATCCACTT 1800
Db 1862 TCACCAATTTCAGCACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAATCCACTT 1921
Qy 1801 GCATCGGATCTCAGGACCTGGCTTTTAGAGCCCTTGGAGCACACCAATACTCTGTGGC 1860
Db 1922 GCATCGGATCTCAGGACCTGGCTTTTAGAGCCCTTGGAGCACACCAATACTCTGTGGC 1981
Qy 1861 GGCACCTGCAGAAACCCAGACACTGGGGAGCTGGTTCCAAAGCTGCCAAGTGTGCA 1920
Db 1982 GGCACCTGCAGAAACCCAGACACTGGGGAGCTGGTTCCAAAGCTGCCAAGTGTGCA 2041
Qy 1921 CTGAGCCCACTTTGGTCAGAGATCGAGGAGATTTGAGAGCGTGGCTTCGGTGGCAACCG 1980
Db 2042 CTGAGCCCACTTTGGTCAGAGATCGAGGAGATTTGAGAGCGTGGCTTCGGTGGCAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCCGCTGAACCTTGGACTAA 2019
```

```
Db 2102 TTGAAGAGAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2140
RESULT 14
MVU34256
LOCUS
DEFINITION
Mice minute virus DNA.
ACCESSION
U34256
VERSION
U34256.1 GI:1464795
KEYWORDS
Mice minute virus
SOURCE
ORGANISM
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 4761)
AUTHORS
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.D., Hook,R.R. Jr. and Riley,L.K.
TITLE
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE
96201434
PUBMED
8609486
REFERENCE
2 (bases 1 to 4761)
AUTHORS
Besselsen,D.G.
TITLE
Direct Submission
SUBMITTED (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
Location/Qualifiers
source
1..4761
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:10794"
ORIGIN
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGAAAAA 60
Db 122 ATGGCTGGAAATGCTTACTCTGATGAGGTTTGGGAGCAACCACTGGTTAAAGAAAAA 181
Qy 61 AGTAACCAAGAAAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAAA 120
Db 182 AGTAACCAAGTATGATTTCTCATTTGTTTAAAAATGAAGATGTTCAATTGAATGAAAA 241
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTACAA 180
Db 242 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTACAA 301
Qy 181 CGAGGAGCGGAACCTACTTGGGACCAACGAGGACATGGAATGGGAACCACTGGAT 240
Db 302 CGAGGAGCGGAACCTACTTGGGACCAACGAGGACATGGAATGGGAATCTTTCAGTGGAT 361
Qy 241 GAAATGACCAAAAAAGCAAGTATTCTTTTGTATTTCTTTGTTTAAAAAATGTTTATTGAA 300
Db 362 GAACTAACCAACAGCAAGTATTCTTTTGTATTTTGTATTTAAAAAGTGTTTATTGAA 421
Qy 301 GTGCTTAAACAAAGAATATATTTCTGCTGATGTTAATTTGTTTGTGCAACATGAATGG 360
Db 422 GTGCTAAGTACAAAGAAACATAGCTCTCTAGTGTATTTAATTTGTTATGTCAGCATGAATGG 481
Qy 361 GGAAGAGCCAGGCTGGCACTGCCATGCTACTTAATTTGGAGGAAAGCACTTTTAGTCAAGCT 420
Db 482 GGAAGAGCCAGGCTGGCACTGCCATGCTACTTAATTTGGAGGAAAGCACTTTTAGTCAAGCT 541
Qy 421 CAAGGAAATGCTGGAGAGGCAACTAAATGTTTACTTGAGCAGATGTTGTTGTAACAGCC 480
Db 542 CAAGGAAATGCTGGAGAGGCACTAAATGTTTACTTGAGCAGATGTTGTTGTAACAGCC 601
Qy 481 TGTAAATGTGCAACTAACACCACTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
```

602	Db	TGCAGTGTGCACTATCACCACGCCGAAGAATAAAGCTGAGAGAAATAGCGGAAGACCAA	661
541	Qy	GAGTGGGTTACTCTACTTATTAAAGCATAAAGCAAAACCAAAAAGACTATACCAAGTG	600
662	Db	GAATGGGTCACTTGGCTTACTTATAGCATAGCAAAACCAAAAAGACTATACCAAATGT	721
601	Qy	GTTCTTTTTGGAAACATGATGTGCTTACTATTTTTTAACCTAAAAAGAAAAAAGCACTAGT	660
722	Db	GTTTGGCTTTGGAAATATGATGCTTACTACTTTTTTAACCAAGAAGAAAAATATGCACTAGT	781
661	Qy	CCACCAAGACGCGAGGCTATTTCTTTAGCAGTGACTCTGGCTGGAAAACTAACCTTTTTTA	720
782	Db	CCACCAAGGACGCGAGGCTATTTCTTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTTTA	841
721	Qy	AAGAAGCGCAGCGCCACTAGTCAGCAAACTATPACACTGATGACATGCGCGCAGAAACG	780
842	Db	AAAGAAGCGCAACGCCATCTAGTGAGCAAACTATACACTGATGNACTGCGCGCAGAAACG	901
781	Qy	GTTGAAACCAAGTAACCACTGCGCAGGAAACTAAGCGCGCGAGAAATTCAAACTTAAAAA	840
902	Db	GTTGAAACCAAGTAACCACTGCGCAGGAAACTAAGCGCGCGAGAAATTCAAACTTAAAAA	961
841	Qy	GAAGTTTCTATTAACAATACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCACACGAG	900
962	Db	GAGGTTTCTATTAACAACCACTTAAAGAGCTGGTGCAATAAAGAGTAACTCACACGAG	1021
901	Qy	GACTTGAATGATGACGACGACAGATTACATTTGAAATGATGGCTCAACACAGGTGGAGAA	960
1022	Db	GACTTGAATGATGACGACGACAGATTACATTTGAAATGATGGCTCAGCAGGTGGAGAA	1081
961	Qy	AACCTGCTGAAAAATACGCTAGAGATTGTACATAACTCTAGCCAGAACCAAAAACAGCA	1020
1082	Db	AACCTGCTTAAAAATACGCTAGAGATCTGTACGCTAACTCTAGCTAGAACCAAAAACAGCC	1141
1021	Qy	TTTGACTTAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCACTGCGCTGAC	1080
1142	Db	TTTGACTTGAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCACTGCGGAC	1201
1081	Qy	ACAAGACCTGACAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCAATGT	1140
1202	Db	ACAAGAACCTGTAGATTTTTGTCTTTTCATGGCTGGAACTACATTAAGTTTGGCATGCT	1261
1141	Qy	ATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTATTTCAATGACCA	1200
1262	Db	ATTTGCTGTGTTCTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTATTTCAATGACCA	1321
1201	Qy	GCACGACACGCAAAATCTATTATGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT	1260
1322	Db	GCCAGTACAGCAAAATCCATCAITTGCAACAGCCATAGCACAGGCAGTTGGTAATGTTGGT	1381
1261	Qy	TGCTATAATCGACCAATGTAACTTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT	1320
1382	Db	TGCTATAATGCAGCAATGTGAACCTTTTCCATTTCAATGACTGTACCAACAAAAACCTGATT	1441
1321	Qy	TGGGTAGAAGAGCTGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT	1380
1442	Db	TGGGTGAAGAGCTGTAACTTTTGGACAGCAAGTAAACCAATTTAAAGCCATTTGCTCT	1501
1381	Qy	GGTCAAACTATTCGCAATGATCAAAAAAGGAAAGGCAAGCAACAGATTGAACCAACCA	1440
1502	Db	GGTCAAACTATTCGCAATGATCAAAAAAGGAAAGGCAAGCAAGTGAACCAACCAACCA	1561
1441	Qy	GTCAATCATGACCACAAATGAGCAATTTACAGTGGTCAAGATAGGCTCGGAGAAAGACCA	1500
1562	Db	GTCAATCATGACCACAAATGAGCAATTTACAGTGGTCAAGATAGGCTCGGAGAGAGACCA	1621
1501	Qy	GAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTTAACACATATCTTGCCT	1560
1622	Db	GAACACACTCAACCAATTAGNACAGAAATGCTCAACATTCATCTAACACATATATTTGCCT	1681
1561	Qy	GGTGACTTTGGTTTTGGTTGACAAAAATGAATGGCCCCATGATTTGTGCTGGTTGGTAAAG	1620

Db	1682	GGTGAC	TTTGGTTGGTTGTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1740
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG		1680
Db	1742	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTGATTGG		1801
Qy	1681	TCAGAAATCTGGGCGGAGCCAAAGTGCGCAACTCTCTATATAATTTACTAGGTTGGGACGC	1740	
Db	1802	TCAGAAATCTGGGCGGAGCCAAAGTGCGCAACTCTCTATATAATTTACTAGGTTGGGACGC		1861
Qy	1741	TCACCATTCAGACACCGGAAAGTACGCCTCTCAGCCAGAACATATGACCTAACTCCACATT	1800	
Db	1862	TCACCATTCAGACACCGGAAAGTACGCCTCTCAGCCAGAACATATGACCTAACTCCACATT		1921
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACAAATACTCTCTGTTCG	1860	
Db	1922	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACAAATACTCTCTGTTCG		1981
Qy	1861	GGCACTCGAATACCCAGAACACCTGGGGAAGCTGGTTTCAAAGCTGCCAAGATGGTCAA	1920	
Db	1982	GGCACTCGAATACCCAGAACACCTGGGGAAGCTGGTTTCAAAGCTGCCAAGATGGTCAA		2041
Qy	1921	CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGGTGCGRACCG	1980	
Db	2042	CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGGTGCGRACCG		2101
Qy	1981	TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGGACTAA	2019	
Db	2102	TTGAGAAAGACTTCAGCGAGCCGCTGAACCTTGGACTAA		2140
RESULT 15				
LOCUS	CO786765	5121 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO2004018689.			PAT 24-MAR-2004
ACCESSION	CO786765			
VERSION	CO786765.1	GI:45721778		
KEYWORDS	synthetic construct			
SOURCE	artificial construct			
ORGANISM	artificial sequences.			
REFERENCE	1			
AUTHORS	Iggo,R. and Malerba,M.			
TITLE	Anti-neoplastic viruses			
JOURNAL	Patent: WO 2004018689-A 1 04-MAR-2004;			
	BTG INTERNATIONAL LIMITED (GB)			
FEATURES	Location/Qualifiers			
source	1. 5121			
	/organism="synthetic construct"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32630"			
	/note="Parvovirus H1(WVM (pH1) - Parvovirus H1 with promoter P4 and left t hairpin from WVM"			
ORIGIN				
Query Match	87.7%;	Score 1771;	DB 6;	Length 5121;
Best Local Similarity	92.3%;	Pred. No. 0;		
Matches 1864;	Conservative 0;	Mismatches 155;	Indels 0;	Gaps 0;
Qy	1	ATGCGTGGAAATGCTTACTCTGTATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	60	
Db	282	ATGCGTGGAAATGCTTACTCTGTATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	341	
Qy	61	AGTAACCGAAGTGTTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGAATGGAAAA	120	
Db	342	AGTAACCGAAGTGTTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGAATGGAAAA	401	
Qy	121	GATATCGATGGGAATAGTTTACAAAAAAGAGCTCGAGGAGGACGAGCTGAAATCTTTACAA	180	
Db	402	GATATCGATGGGAATAGTTTACAAAAAAGAGCTCGAGGAGGACGAGCTGAAATCTTTACAA	461	
Qy	181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGGACATGGAAATGGGAAACCAAGTGGAT	240	

Db 462 CGAGGAGCGGAACTACTTGGGACAAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 521  
Qy 241 GAAATGACCAAAAAGCAAGTATTCATTTTGGTATCTTTGGTTAAAAAATGTTTATTTGAA 300  
Db 522 GAAATGACCAAAAAGCAAGTATTCATTTTGGTATCTTTGGTAAAAAATGTTTATTTGAA 581  
Qy 301 GTGCTTAAACCAAAAGAAATATATTTTCTGGTGATGTTAAATTTGGTTGTGCAACATGAATGG 360  
Db 582 GTGCTTAAACCAAAAGAAATATATTTTCTGGTGATGTTAAATTTGGTTGTGCAACATGAATGG 641  
Qy 361 GGA AAAAGACCAAGCTGGCACTGCGCATGTACTTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420  
Db 642 GGA AAAAGACCAAGCTGGCACTGCGCATGTACTTAATTTGGAGGAAAGGACTTTAGTCAAGCT 701  
Qy 421 CAAGGGAATTTGGTGGAGAGCACTAAATTTTACTTGGAGCAGATGTTGGTAAACAGCC 480  
Db 702 CAAGGGAATTTGGTGGAGAGCACTAAATTTTACTTGGAGCAGATGTTGGTAAACAGCC 761  
Qy 481 TGTAAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAAT 540  
Db 762 TGTAAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAAT 821  
Qy 541 GAGTGGGTTACTTACTTACTTATTAAGCATTAAGCAACCAAAAAGACTATACCAAGTGT 600  
Db 822 GAGTGGGTTACTTACTTACTTATTAAGCATTAAGCAACCAAAAAGACTATACCAAGTGT 881  
Qy 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATAAAGCACTAGT 660  
Db 882 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATAAAGCACTAGT 941  
Qy 661 CCACCAAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACCTAACTTTTAA 720  
Db 942 CCACCAAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACCTAACTTTTAA 1001  
Qy 721 AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG 780  
Db 1002 AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGATGAACCAAGAAACG 1061  
Qy 781 GTTGA AACCAACAGTAACCACTGGCGAGAAACTAAGCGCGGAGAAATTAACCTAA AAAA 840  
Db 1062 GTCGAGACCAAGTACCACTGCACAGGAAGCTTAGCGCGGAGAAATTAACCTAAGAG 1121  
Qy 841 GAAGTTTCTAATTAACCTACCTTAAGAGCTGGTGATTAAGAGTAACCTCACAGAG 900  
Db 1122 GAGGCTCTCGAATTAACCACTCAAGAGCTGGTACATAAAGAGTAACCTCACAGAA 1181  
Qy 901 GACTGGATGATGACGACGACGTTACATTTGAATGATGCTCAACCAAGTGGAGAA 960  
Db 1182 GACTGGATGATGACGACGACGTTACATTTGAATGATGCTCAACCAAGTGGAGAA 1241  
Qy 961 AACCTGTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAAACCAAAACAGCA 1020  
Db 1242 AACCTGTGAAAAATACCTAGAGATCTGTACACTGACTAGCAGAGAACCAAAACAGCC 1301  
Qy 1021 TTTGACTTAATTTAGAAAAGCTGAAAACAGCAAACTAACCTTTTCACTGCTGAC 1080  
Db 1302 TTTGACTTGAATTTGAAAAAGCTGAAAACAGCAAACTAGCCAACTTTTCCATGGCTAGC 1361  
Qy 1081 ACAAGAGCTGCAAGATTTTGGCTTTTCACTGGCTGGAATGATGTTAAAGTTTGGCAATGCT 1140  
Db 1362 ACCAGAACCTGTGAATCTTTGTGTAGCATGGCTGGAACTATATTAAGTCTGCCATGCC 1421  
Qy 1141 ATTTGCTGTGTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTATTTTATGACCA 1200  
Db 1422 ATCTGTGTGTGTAATAGACAAGGAGGCAAAAGGAACACTGTGCTCTTTCACGGACCA 1481  
Qy 1201 GCAGACAGGCAAACTATTTATGCAAGGCCATAGCAGCAAGCTTTGGCAATGTTGGT 1260  
Db 1482 GCCAGACAGGCAAACTATTTATGCAAGGCCATAGCAGCAAGCTTTGGTAAATGTTGGT 1541  
Qy 1261 TGCCTATAATGCAGCAATGTAACTTTTCCATTTTAAATGACTGTATCAACCAAGAACTTGATT 1320  
Db 1542 TGTATACATGCTGCCAATGTGAACCTTTCCATTTTAAATGACTGTATCAACCAAACTTGATT 1601

Qy 1321 TGGGTAGAAGAAGCTCGTTAACTTTTGACAGCAAGTAATAACAGTTTAAAGCATTTCCTCT 1380  
Db 1602 TGGGTGGAAGAAGCTCGTTAACTTTTGCCAGCAAGTAATAACAAATTCAAAGCTATTTGTTCT 1661  
Qy 1381 GGTCAAACTATTTGCGATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440  
Db 1662 GGCCAAACTATGCGATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1721  
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGATAGGTGCGCAAGAAAGACCA 1500  
Db 1722 GTTATTTATGACCACCAACGAGAACATTACCGTGGTTAGAAATAGGCTGTGAGGAAAGACCA 1781  
Qy 1501 GAAACACATCAACCAATCAGAGACAGAAATGCTTTAACTTCACTTAACACATCTTGGCT 1560  
Db 1782 GAAACACATCAACCAATCAGAGACAGAAATGCTTCAACATTCACCTGACAGTACACTACCT 1841  
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAAATGAATGGCCCATGATTTGTGCTTTGGTTGTAAG 1620  
Db 1842 GGTGACTTTGGTTGGTTGATTAAGCACGAATGGCCTCTGATCTGTGCTTTGGTTGGTGAAG 1901  
Qy 1621 AATGGTTTCAATCTTACCATGCGCAAGCTACTGTCTTAATGGGGCAAAAGTTCTCTGATTGG 1680  
Db 1902 AATGGTTTCAATCTTACCATGCGCTTTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1961  
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCCTTATAAAATTTACTAGGTTTCGGCAGCG 1740  
Db 1962 TCAGAGGACTGGCGGAGCCGAAGCTAGACACTCCTATAAAATTCGCTAGGTTCAATGCGC 2021  
Qy 1741 TCACCAATTCACGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 2022 TCACCAATCTCTGACTCCGAGAAAGTAGCGCTCTCAGCCAAAACTAGCGCTTTACTCCACTT 2081  
Qy 1801 GCATCGGATCTCGAGACCTGGCTTTAGAGCCTTGGAGCACACCAAACTACTCTCTGTTGCG 1860  
Db 2082 GCATCGGACCTTGGGAGCCTAGCTCTAGAGCCTTGGAGCACACCAAACTACTCTCTGTTGCG 2141  
Qy 1861 GGCACCTGCAGAAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Db 2142 GGCACCTGCAGCAAGCCAAAACACTGGGAGGCTGGTTCCACAGCCTGCCAAGGTGCTCAA 2201  
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTTCGGAAACCG 1980  
Db 2202 CGGAGCCCAACTGGTTCGAGATCGAGGCGGATTTGAGAGCTTGTCTTCAGTCAAGAAACAG 2261  
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACCTTGGACTAA 2019  
Db 2262 TTGGAGAGCGACTTCAACGAGAGCTGACCTTGGACTAA 2300

Search completed: January 22, 2005, 21:19:21

Job time : 8547.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds

(without alignments)  
11150.532 Million cell updates/sec

Title: US-10-069-056-8

Perfect score: 2019

Sequence: 1 atggctggaatgcttactc.....agccgcgtgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2019	100.0	2019	5	Ad02801 Parvoviru
2	2017.4	99.9	2019	5	Ad02797 Parvoviru
3	2017.4	99.9	5149	12	Adg39767 Minute v1
4	2015.8	99.8	2019	5	Ad02805 Parvoviru
5	2015.8	99.8	2019	5	Ad02803 Parvoviru
6	2015.8	99.8	2019	5	Ad02799 Parvoviru
7	1040.4	51.5	5049	2	Aat15311 Non-atten
8	1040.4	51.5	5049	2	Aat15312 Attenuate
9	1040.4	51.5	5049	2	Aat88324 Attenuate
10	1038.8	51.5	5049	2	Aat88321 Attenuate
11	1038.8	51.5	5049	2	Aat88320 Canine pa
12	901	44.6	3524	1	Aan40252 Sequence
13	307.4	15.2	374	6	Abq95626 Tumour su
14	285.8	14.2	421	6	Abq94779 Tumour su
15	284.8	14.1	423	6	Abq94724 Tumour su
16	284.8	14.1	464	6	Abq94719 Tumour su
17	284.8	14.1	473	6	Abq94716 Tumour su
18	284.8	14.1	486	6	Abq94707 Tumour su
19	284.8	14.1	491	6	Abq94777 Tumour su
20	284.2	14.1	420	6	Abq94787 Tumour su
21	283.8	14.1	343	6	Abq94737 Tumour su

22	283.8	14.1	343	6	AbQ94739	Tumour su
23	283.8	14.1	403	6	AbQ94782	Tumour su
24	283.2	14.0	497	6	AbQ94718	Tumour su
25	280.4	13.9	340	6	AbQ94732	Tumour su
26	278.2	13.8	342	6	AbQ94765	Tumour su
27	273.8	13.6	420	6	AbQ94781	Tumour su
28	269.6	13.4	324	6	AbQ94740	Tumour su
29	263.4	13.0	339	6	AbQ95625	Tumour su
30	261.6	13.0	424	6	AbQ94778	Tumour su
31	163.8	8.1	468	6	AbQ94873	Tumour su
32	162.2	8.0	451	6	AbQ94868	Tumour su
33	162.2	8.0	457	6	AbQ94867	Tumour su
34	162.2	8.0	465	6	AbQ94874	Tumour su
35	162.2	8.0	469	6	AbQ94869	Tumour su
36	162.2	8.0	472	6	AbQ94880	Tumour su
37	162.2	8.0	515	6	AbQ94793	Tumour su
38	162.2	8.0	516	6	AbQ94807	Tumour su
39	162.2	8.0	516	6	AbQ94794	Tumour su
40	162.2	8.0	516	6	AbQ94800	Tumour su
41	162.2	8.0	517	6	AbQ94811	Tumour su
42	162.2	8.0	530	6	AbQ94791	Tumour su
43	162.2	8.0	530	6	AbQ94799	Tumour su
44	161	8.0	516	6	AbQ94796	Tumour su
45	160.6	8.0	530	6	AbQ94805	Tumour su

ALIGNMENTS

RESULT 1

AD02801

ID AD02801 standard; DNA; 2019 BP.

XX AC AD02801;

XX DT 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX DE Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.

KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.  
XX OS Parvovirus.  
XX OS Synthetic.  
FH Key Location/Qualifiers  
FT CDS 1..2019  
FT /tag= a  
FT /product= "NS1 variant (T363A) protein"  
FT mutation replace(1187, A)  
FT /tag= b

XX EPI077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nuesch J, Rommelaere J;

XX WFI; 2001-212717/22.

XX P-PSDB; AAY72706.

XX Novel parvovirus non-structure protein variant, useful for treating  
XX tumoral diseases, has a shifted equilibrium between DNA replication and  
XX transcription activities, and cytotoxic activity.

XX Claim 7; Page 16-19; 41pp; English.



XX	The present sequence is a DNA encoding parvovirus non-structure protein 1 (NS1) variant (7363A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the cytotoxicity activity. These variants are useful as toxins for treating tumoural diseases. The variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-2003 to correct OS field.)									
XX	SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;									
	Query Match	100.0%;	Score 2019;	DB 5;	Length 2019;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 2019;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
Qy	1	ATGGCTGGAATGCTTACTCTGATGAAGTCTTGGGACCAACCACTGTTTAAAGGAAAA	60							
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTCTTGGGACCAACCACTGTTTAAAGGAAAA	60							
Qy	61	AGTAACGAGGAAGTCTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGGAATGAAAA	120							
Db	61	AGTAACGAGGAAGTCTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGGAATGAAAA	120							
Qy	121	GATATCGGATGGATAGTTACAAAGAGCTGCGAGGACGAGCTGGAATCTTTACAA	180							
Db	121	GATATCGGATGGATAGTTACAAAGAGCTGCGAGGACGAGCTGGAATCTTTACAA	180							
Qy	181	CGAGGACGGAATCTCTGGGACCAAGCGAGCATGGAATGGGAAACACAGTGGAT	240							
Db	181	CGAGGACGGAATCTCTGGGACCAAGCGAGCATGGAATGGGAAACACAGTGGAT	240							
Qy	241	GAATGACCAAAAGCAAGTATTCATTTTGTCTTTGTTTAAATAAGTATTTATTTGAA	300							
Db	241	GAATGACCAAAAGCAAGTATTCATTTTGTCTTTGTTTAAATAAGTATTTATTTGAA	300							
Qy	301	GTGCTTAACCAAGATATATTCCTGGTGATGTTAAATGGTTTGTGCAACATGATGG	360							
Db	301	GTGCTTAACCAAGATATATTCCTGGTGATGTTAAATGGTTTGTGCAACATGATGG	360							
Qy	361	GGAAAGACCAAGGCTGGCACTGCCATGTAATTTGGAGGAAGGACTTTAGTCAAGCT	420							
Db	361	GGAAAGACCAAGGCTGGCACTGCCATGTAATTTGGAGGAAGGACTTTAGTCAAGCT	420							
Qy	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACGCC	480							
Db	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACGCC	480							
Qy	481	TGTAATGTGCAACTACACACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540							
Db	481	TGTAATGTGCAACTACACACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540							
Qy	541	GAGTGGGTACTCTTACTTATTAAGCATAAGCAACCAACCAAGACTATACCAAGTGT	600							
Db	541	GAGTGGGTACTCTTACTTATTAAGCATAAGCAACCAACCAAGACTATACCAAGTGT	600							
Qy	601	GTTCTTTTGGAAACATGATGTTTACTATTTTAACTAAAGAGAAATTAAGCACTAGT	660							
Db	601	GTTCTTTTGGAAACATGATGTTTACTATTTTAACTAAAGAGAAATTAAGCACTAGT	660							
Qy	661	CCACCAAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTACTTTTAA	720							
Db	661	CCACCAAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTACTTTTAA	720							
Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780							
Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780							
Qy	781	GTTGAAACCAACAGTAACCACTGGCGAGAACTAAGCGCGGAGAAATCAAACTAAAAA	840							
Db	781	GTTGAAACCAACAGTAACCACTGGCGAGAACTAAGCGCGGAGAAATCAAACTAAAAA	840							
Qy	841	GAAGTTTCTATTAAACTACATTAAAGACTGGTGCATTAAGAGCTTACCTCACCAGAG	900							
Db										

841	GAAGTTTCTATTAAAACTACATTAAAGAGCTGGTGCATATAAAGAGTAACTCCTCACCAGAG	900
901	GACTGATGATGATGAGCAGACAGACAGTACATTGAATGATGCTCAACACAGGTGAGAA	960
901	GACTGATGATGATGAGCAGACAGACAGTACATTGAATGATGCTCAACACAGGTGAGAA	960
961	AACCTGCTGAAAAAATACGCTAGAGATTGTACATNACTCTAGCCAGAAACCAAAACAGCA	1020
961	AACCTGCTGAAAAAATACGCTAGAGATTGTACATNACTCTAGCCAGAAACCAAAACAGCA	1020
1021	TTTGACTTAATTTTAGAAAAAGCTGAACCAACCAACTAACCACTTTTCACTGCTCTGAC	1080
1021	TTTGACTTAATTTTAGAAAAAGCTGAACCAACCAACTAACCACTTTTCACTGCTCTGAC	1080
1081	ACAAGAGCCTGCAGAAATTTTGTCTTTCATGCTGGAATATATGTTTAAAGTTTGCCTATGCT	1140
1081	ACAAGAGCCTGCAGAAATTTTGTCTTTCATGCTGGAATATATGTTTAAAGTTTGCCTATGCT	1140
1141	ATTGCTGTGTTTTTAAACAGAGAGGCAAAAGAAATACCTGTTTTTATTCATGAGCA	1200
1141	ATTGCTGTGTTTTTAAACAGAGAGGCAAAAGAAATACCTGTTTTTATTCATGAGCA	1200
1201	GCAGCAGAGGCAAAATCTATTATTCACAAGCCATAGCACAAGCAGTTGGCAATGTTGT	1260
1201	GCAGCAGAGGCAAAATCTATTATTCACAAGCCATAGCACAAGCAGTTGGCAATGTTGT	1260
1261	TGCTATAATGCGACCAATGTAACTTTCCATTTAATGACTGTACCAACCAAGAACTTTGAT	1320
1261	TGCTATAATGCGACCAATGTAACTTTCCATTTAATGACTGTACCAACCAAGAACTTTGAT	1320
1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCACTTTTAAAGCCATTTGCTCT	1380
1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCACTTTTAAAGCCATTTGCTCT	1380
1381	GGTCAAACTATTTCGCAATTCATCAAAAGGAAAGGAGCAGACAGATTTGAAACCAACCA	1440
1381	GGTCAAACTATTTCGCAATTCATCAAAAGGAAAGGAGCAGACAGATTTGAAACCAACCA	1440
1441	GTCTATCATGACCAAAATGAGAACTTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA	1500
1441	GTCTATCATGACCAAAATGAGAACTTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA	1500
1501	GAAACACTCAACCAATCAGACAGAGATGCTTAACTATCATCTAACACATACCTTGCCT	1560
1501	GAAACACTCAACCAATCAGACAGAGATGCTTAACTATCATCTAACACATACCTTGCCT	1560
1561	GGTGACTTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTTGTGGTTGTTAAAG	1620
1561	GGTGACTTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTTGTGGTTGTTAAAG	1620
1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTTCTGATTGG	1680
1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTTCTGATTGG	1680
1681	TCAGAAAACCTGGCGAGCCAAAGGTGCCAATCCTATATAATTTACTAGTTTGGCACCG	1740
1681	TCAGAAAACCTGGCGAGCCAAAGGTGCCAATCCTATATAATTTACTAGTTTGGCACCG	1740
1741	TCACCAATTCACCAACCGAAAGTACGCTCTCAGCAGAACTATGCACTTAATCCACTT	1800
1741	TCACCAATTCACCAACCGAAAGTACGCTCTCAGCAGAACTATGCACTTAATCCACTT	1800
1801	GCATCGGATCTCGAGGACCTGCTTTTAGAGCCTTTGGAGCACCAAAATACTCTGTGCG	1860
1801	GCATCGGATCTCGAGGACCTGCTTTTAGAGCCTTTGGAGCACCAAAATACTCTGTGCG	1860
1861	GGAATGCGAGAAACCCAGAACTGCGGAGAGCTGTTTCCAAAGCCTGCCAAGATGTTCAA	1920
1861	GGAATGCGAGAAACCCAGAACTGCGGAGAGCTGTTTCCAAAGCCTGCCAAGATGTTCAA	1920
1921	CTGAGCCCAACTTGTGTGAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGTGGGAAACCG	1980
1921	CTGAGCCCAACTTGTGTGAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGTGGGAAACCG	1980



QY	1981	TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA	2019
Db	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA	2019
RESULT 2			
AAD02797			
ID	AAD02797 standard; DNA; 2019 BP.		
XX			
AC	AAD02797;		
XX			
DT	06-AUG-2003 (revised)		
DT	31-MAY-2001 (first entry)		
XX			
DE	Parvovirus non-structure protein 1 (NS1) wild-type DNA.		
XX			
KW	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;		
KW	tumoural disease; gene therapy; ds.		
XX			
OS	Parvovirus.		
XX			
FH	Location/Qualifiers		
FT	1. .2019		
FT	/*tag= a		
FT	/product= "Parvovirus NS1 protein"		
XX			
PN	EP1077260-A1.		
XX			
PD	21-FEB-2001.		
XX			
PF	13-AUG-1999; 99BP-00115161.		
XX			
PR	13-AUG-1999; 99BP-00115161.		
XX			
PA	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
XX			
PI	Nuesch J, Rommelaere J;		
XX			
DR	WPI; 2001-212717/22.		
DR	P-PSDB; AAY72702.		
XX			
PT	Novel parvovirus non-structure protein variant, useful for treating		
PT	tumoral diseases, has a shifted equilibrium between DNA replication and		
PT	transcription activities, and cytotoxic activity.		
XX			
PS	Disclosure; Fig 1; 41pp; English.		
XX			
CC	The present sequence is a wild type DNA encoding parvovirus non-		
CC	structure protein 1 (NS1). The present invention relates to the variants		
CC	of the parvovirus non-structure protein (NS1) having a shifted		
CC	equilibrium between the DNA replication and transcription activities, and		
CC	the cytotoxicity activity. These variants are useful as toxins for		
CC	treating tumoural diseases. The variant DNAs are useful as vectors for		
CC	gene therapy. (Updated on 06-AUG-2003 to correct OS field.)		
XX			
SQ	Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;		
Query Match 99.9%; Score 2017.4; DB 5; Length 2019;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
QY	61	AGTAACCAAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
Db	61	AGTAACCAAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA	180
Db	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA	180

```
QY 1261 TGCTATTAATGAGCAATGTAACCTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320
D 1261 TGCTATTAATGAGCAATGTAACCTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320
QY 1321 TGGGTAGAAGAGCTGTAACTTTGGACAGCAAGTAACAGTTTAAAGCCATTGCTCT 1380
D 1321 TGGGTAGAAGAGCTGTAACTTTGGACAGCAAGTAACAGTTTAAAGCCATTGCTCT 1380
QY 1381 GGTCAAACTATTCCGATTGATCAAAAAGGAAAGGAGCAGCAAAACAGATTGAAACCAACCA 1440
D 1381 GGTCAAACTATTCCGATTGATCAAAAAGGAAAGGAGCAGCAAAACAGATTGAAACCAACCA 1440
QY 1441 GTCATCATGACCAAAATGAAACATTACAGTGTGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
D 1441 GTCATCATGACCAAAATGAAACATTACAGTGTGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGACAGACAGAAATGCTTAACATTCATCTAACACATACCTTGCCT 1560
D 1501 GAACACACTCAACCAATCAGACAGACAGAAATGCTTAACATTCATCTAACACATACCTTGCCT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
D 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
D 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAACTGGGGGAGCCAAAGGTGCCAATCTCTATAATTTACTAGTTGGGACCG 1740
D 1681 TCAGAAACTGGGGGAGCCAAAGGTGCCAATCTCTATAATTTACTAGTTGGGACCG 1740
QY 1741 TCACCATTCAGACACCCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAATCTCCACTT 1800
D 1741 TCACCATTCAGACACCCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAATCTCCACTT 1800
QY 1801 GCATCGAATTCGAGGACCTGGCTTTAGAGCTTTGAGGACACACCAATATCTCTGTGG 1860
D 1801 GCATCGAATTCGAGGACCTGGCTTTAGAGCTTTGAGGACACACCAATATCTCTGTGG 1860
QY 1861 GGCACCTGCAGAAACCCAGACACTGGGAGCTGTTCCAAAGCCTGCCAAGATGCTCAA 1920
D 1861 GGCACCTGCAGAAACCCAGACACTGGGAGCTGTTCCAAAGCCTGCCAAGATGCTCAA 1920
QY 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
D 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
QY 1981 TTGAAGAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
D 1981 TTGAAGAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
```

## RESULT 3

```
ADG39767
ID ADG39767 standard; DNA; 5149 BP.
XX
AC AC
ADG39767;
XX
DT 11-MAR-2004 (first entry)
XX
DE Minute virus from mouse genomic DNA sequence SEQ ID NO:29.
XX
KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
OS Mice minute virus.
XX
PN WO2003104392-A2.
XX
PD 18-DEC-2003.
```

```
XX 02-DEC-2002; 2002WO-US038423.
PF
XX
PR 18-DEC-2001; 2001US-0341919P.
XX
PA (UYN-) UNIV NORTH CAROLINA.
XX
PI Samulski RJ, Rabinowitz JE;
XX WPI; 2004-062324/06.
XX
PT New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
PS Disclosure; SEQ ID NO 29; 115pp; English.
XX
CC The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC (1); (2) a cell comprising (I), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus; the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 5149 BP; 1718 A; 1045 C; 1124 G; 1262 T; 0 U; 0 Other;
Query Match 99.9%; Score 2017.4; DB 12; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
D 261 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 320
QY 61 AGTAACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
D 321 AGTAACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 380
QY 121 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGGAGGAGCTGAAATCTTTTACAA 180
D 381 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGGAGGAGCTGAAATCTTTTACAA 440
QY 181 CGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCACTGGAT 240
D 441 CGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCACTGGAT 500
QY 241 GAAATGACCAAAAACAGATATTCTTTTGAATCTTTGGTTAAAAAATGTTTATTCAA 300
D 501 GAAATGACCAAAAACAGATATTCTTTTGAATCTTTGGTTAAAAAATGTTTATTCAA 560
QY 301 GTGCTTAACACAAAGATATATTTCTCGTGTATGTTAATTTGTTGTGCAACATCAATGG 360
D 561 GTGCTTAACACAAAGATATATTTCTCGTGTATGTTAATTTGTTGTGCAACATCAATGG 620
```

QY 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGACCTTTAGTCAAGCT 420  
DB 621 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGACCTTTAGTCAAGCT 680  
QY 421 CAAAGGAAATGGTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAAACAGCC 480  
DB 681 CAAAGGAAATGGTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAAACAGCC 740  
QY 481 TGTAAATGTGCAACTAACAACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAGAAAT 540  
DB 741 TGTAAATGTGCAACTAACAACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAGAAAT 800  
QY 541 GAGTGGGTTACTCTACTTATATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600  
DB 801 GAGTGGGTTACTCTACTTATATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 860  
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAAGAAATAGCACTAGT 660  
DB 861 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAAGAAATAGCACTAGT 920  
QY 661 CCAAGAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720  
DB 921 CCAAGAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 980  
QY 721 AAGAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
DB 981 AAGAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 1040  
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTTAAGCGGCGAGAAATCAAACTAAAAAA 840  
DB 1041 GTTGAACACACAGTAACCACTGCGCAGGAAACTTAAGCGGCGAGAAATCAAACTAAAAAA 1100  
QY 841 GAAAGTTTCTATTAAAACTACCTTAAGAGCTGGTGCCATAAAAGAGTAACCTCACAGAG 900  
DB 1101 GAAAGTTTCTATTAAAACTACCTTAAGAGCTGGTGCCATAAAAGAGTAACCTCACAGAG 1160  
QY 901 GACTGATGATGATGACGAGCAGAGTACATTTGAATGATGCTCAACAGGTGGAGAA 960  
DB 1161 GACTGATGATGATGACGAGCAGAGTACATTTGAATGATGCTCAACAGGTGGAGAA 1220  
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAAACCAAAACAGCA 1020  
DB 1221 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAAACCAAAACAGCA 1280  
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080  
DB 1281 TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1340  
QY 1081 ACAAGAGCTGCAAGATTTTGTCTTTTTCATGGCTGGAATATGTTAAAGTTTGCATGCT 1140  
DB 1341 ACAAGAGCTGCAAGATTTTGTCTTTTTCATGGCTGGAATATGTTAAAGTTTGCATGCT 1400  
QY 1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTCATGACCA 1200  
DB 1401 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTCATGACCA 1460  
QY 1201 GCCAGCAGCAGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATTTGGT 1260  
DB 1461 GCCAGCAGCAGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATTTGGT 1520  
QY 1261 TGCTATAATGCGCAATGTAACCTTTCCATTTTAAATGATGCTACCAAGAACTTGATT 1320  
DB 1521 TGCTATAATGCGCAATGTAACCTTTCCATTTTAAATGATGCTACCAAGAACTTGATT 1580  
QY 1321 TGGGTAGAACAGCTGTTACTTTTGGACAGCAAGTAAGCAAGTTTAAAGCACTTTGCTCT 1380  
DB 1581 TGGGTAGAACAGCTGTTACTTTTGGACAGCAAGTAAGCAAGTTTAAAGCACTTTGCTCT 1640  
QY 1381 GGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1440  
DB 1641 GGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1700  
QY 1441 GTCATCATGACCAAAATAGAAACATTACAGTGGTCAGAAATAGGCTGCGGAAGAAAGACCA 1500

DB 1701 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGGAAGAAAGACCA 1760  
QY 1501 GAAACACACTCAACCAATCAGAGACAGAATGCTTAACTCATCTTAACATACCTTGGCT 1560  
DB 1761 GAAACACACTCAACCAATCAGAGACAGAATGCTTAACTCATCTTAACATACCTTGGCT 1820  
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620  
DB 1821 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1880  
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTATTGG 1680  
DB 1881 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTATTGG 1940  
QY 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGTTGCGCACGC 1740  
DB 1941 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGTTGCGCACGC 2000  
QY 1741 TCACCAATTCACGACACCCGAAAGTAGCCTCTCAGCCAGAACTATGCACTTAACCTCCACT 1800  
DB 2001 TCACCAATTCACGACACCCGAAAGTAGCCTCTCAGCCAGAACTATGCACTTAACCTCCACT 2060  
QY 1801 GCATCGATCTCGAGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTTTGGG 1860  
DB 2061 GCATCGATCTCGAGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTTTGGG 2120  
QY 1861 GGCACCTGCAGAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 1920  
DB 2121 GGCACCTGCAGAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 2180  
QY 1921 CTGAGCCCAACTTGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG 1980  
DB 2181 CTGAGCCCAACTTGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG 2240  
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019  
DB 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2279

## RESULT 4

AAD02805

ID AAD02805 standard; DNA; 2019 BP.

XX AAD02805;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

XX Key Location/Qualifiers

FH 1. .2019

FT /tag= a

FT /product= "NS1 variant (T463A) protein"

FT replace (1387, A)

FT /tag= b

FT mutation

FT EP1077260-A1.

PN 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PA

XX Nuéesch J, Rommelaere J;  
PI WPI: 2001-212717/22.  
XX P-PSDB; AAY72710.  
XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.  
XX Claim 7; Page 27-30; 41pp; English.  
XX The present sequence is a DNA encoding parvovirus non-structure protein 1  
CC (NS1) variant (7463A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoral diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;  
Query Match 99.88; Score 2015.8; DB 5; Length 2019;  
Best Local Similarity 99.94; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGGTTAAAGGAAAA 60  
DB 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGGTTAAAGGAAAA 60  
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAATGAAAAA 120  
DB 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAATGAAAAA 120  
QY 121 GATATCGGATGGATAGTATCAAAAAGAGCTGCAGGAGCAGAGCTGAATCTTTACAA 180  
DB 121 GATATCGGATGGATAGTATCAAAAAGAGCTGCAGGAGCAGAGCTGAATCTTTACAA 180  
QY 181 CGAGGAGCGGAACCTACTTGGGACCAAGCGAGCATGGAATGGGAAACCAAGTGGAT 240  
DB 181 CGAGGAGCGGAACCTACTTGGGACCAAGCGAGCATGGAATGGGAAACCAAGTGGAT 240  
QY 241 GAAATGACAAAAGCAAGTATTCATTTTGGATCTTTGGTTAAAAATGTTTATTTGAA 300  
DB 241 GAAATGACAAAAGCAAGTATTCATTTTGGATCTTTGGTTAAAAATGTTTATTTGAA 300  
QY 301 GTGCTTAACCAAGAAATATTTCTGTTGATGTTAAATTTGTTTGTCAACATGAATGG 360  
DB 301 GTGCTTAACCAAGAAATATTTCTGTTGATGTTAAATTTGTTTGTCAACATGAATGG 360  
QY 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420  
DB 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420  
QY 421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTTGTAACAGCC 480  
DB 421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTTGTAACAGCC 480  
QY 481 TGTAAATGTGCAACTAACCACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540  
DB 481 TGTAAATGTGCAACTAACCACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540  
QY 541 GAGTGGGTTACTCTTACTTACTTATAAGCATAAGCAAAACCAAAAAGACTATACCAAGTGT 600  
DB 541 GAGTGGGTTACTCTTACTTACTTATAAGCATAAGCAAAACCAAAAAGACTATACCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAGAAAAATAGCACTAGT 660  
DB 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAGAAAAATAGCACTAGT 660  
QY 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTAA 720  
DB 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTAA 720

QY 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
DB 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTAAGCGCGGAGAAATCAAACTAAAAAA 840  
DB 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTAAGCGCGGAGAAATCAAACTAAAAAA 840  
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900  
DB 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900  
QY 901 GACTGATGATGATGAGCAGCAGACAGTACATTGAAATGATGCTCAACACAGGTGAGAAA 960  
DB 901 GACTGATGATGATGAGCAGCAGACAGTACATTGAAATGATGCTCAACACAGGTGAGAAA 960  
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAACAGCA 1020  
DB 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAACAGCA 1020  
QY 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACAGCAACTAAACCACTTTTCACTGCTGAC 1080  
DB 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACAGCAACTAAACCACTTTTCACTGCTGAC 1080  
QY 1081 ACAAGAGCCTGAGAAATTTTCTTCAATGCTGGAATATGTTTAAAGTTTGCCTATGCT 1140  
DB 1081 ACAAGAGCCTGAGAAATTTTCTTCAATGCTGGAATATGTTTAAAGTTTGCCTATGCT 1140  
QY 1141 ATTGCTGTGTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTCAATGAGCA 1200  
DB 1141 ATTGCTGTGTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTCAATGAGCA 1200  
QY 1201 GCAGACAGGCAAACTATTATTCACAAAGCCATAGCAGACAGCTGGCAATGTTGCT 1260  
DB 1201 GCAGACAGGCAAACTATTATTCACAAAGCCATAGCAGACAGCTGGCAATGTTGCT 1260  
QY 1261 TGTATAATGACGCAAACTGTAATTTCAATTAATGACTGTACCAACAAAGAACTTGAAT 1320  
DB 1261 TGTATAATGACGCAAACTGTAATTTCAATTAATGACTGTACCAACAAAGAACTTGAAT 1320  
QY 1321 TGGGTAGAAGAGCTGGTAACTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
DB 1321 TGGGTAGAAGAGCTGGTAACTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
QY 1381 GGTCAAACTATTTCGCAATGATCAAAAAGGAAAGCAGCAACAGATTGAAACCAACCA 1440  
DB 1381 GGTCAAACTATTTCGCAATGATCAAAAAGGAAAGCAGCAACAGATTGAAACCAACCA 1440  
QY 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500  
DB 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500  
QY 1501 GAACACATCAACCAATCAGAGACAGATGCTTAACATTCATCTAACACATACCTTGCCT 1560  
DB 1501 GAACACATCAACCAATCAGAGACAGATGCTTAACATTCATCTAACACATACCTTGCCT 1560  
QY 1561 GGTGACTTTGGTTTGGTTGACAAAAATGAAATGGCCCATGATTGCTTGGTTGGTAAAG 1620  
DB 1561 GGTGACTTTGGTTTGGTTGACAAAAATGAAATGGCCCATGATTGCTTGGTTGGTAAAG 1620  
QY 1621 AATGTTTACCAATCTACCATGCGCAAGTACTGTGCTAAATGGGCAAAAGTTCTCTGATTG 1680  
DB 1621 AATGTTTACCAATCTACCATGCGCAAGTACTGTGCTAAATGGGCAAAAGTTCTCTGATTG 1680  
QY 1681 TCAGAAAACTGGCGGAGCCAAAGTGGCAACTCTCTATAAATTTTACTAGTTTCGCGACGC 1740  
DB 1681 TCAGAAAACTGGCGGAGCCAAAGTGGCAACTCTCTATAAATTTTACTAGTTTCGCGACGC 1740  
QY 1741 TCACCAATTCAGACACCGGAAAGTACGCTCTCAGCCAGACTATGCACTAATCCACTT 1800  
DB 1741 TCACCAATTCAGACACCGGAAAGTACGCTCTCAGCCAGACTATGCACTAATCCACTT 1800



Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACGAGCAAACTAACCAACTTTTTCACCTGCCTGAC 1080  
QY 1081 ACAAGAGCCTGCAAGATTTTGTCTTTTCATGGCTGGAACCTATCTATAAGTTTCCCATGCT 1140  
Db 1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGGCTGGAACCTATCTATAAGTTTCCCATGCT 1140  
QY 1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAGAAATACTGTTTATTTATTCATGACCA 1200  
Db 1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAGAAATACTGTTTATTTATTCATGACCA 1200  
QY 1201 GCCAGCAGGCAAACTTATTTATTCAGACGCTATAGCAAGCAGTGGCAATGTTGCT 1260  
Db 1201 GCCAGCAGGCAAACTTATTTATTCAGACGCTATAGCAAGCAGTGGCAATGTTGCT 1260  
QY 1261 TGCTATAATCAGCAGTAACTTTTCCATTTAATGACTGTACCAAGCAAGCACTTGAAT 1320  
Db 1261 TGCTATAATCAGCAGTAACTTTTCCATTTAATGACTGTACCAAGCAAGCACTTGAAT 1320  
QY 1321 TGCTTAGAAGAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTTAAAGCAATTTGCTCT 1380  
Db 1321 TGCTTAGAAGAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTTAAAGCAATTTGCTCT 1380  
QY 1381 GGTCAAACTATTGCGATTGATCAAAAGGAAAGGAGGAGCAAGCAAGTGAACCAACCA 1440  
Db 1381 GGTCAAACTATTGCGATTGATCAAAAGGAAAGGAGGAGCAAGCAAGTGAACCAACCA 1440  
QY 1441 GTCATCATGACCAATGAGACATTACAGTGTCTGAGATAGGCTGCGAAGAAAGACCA 1500  
Db 1441 GTCATCATGACCAATGAGACATTACAGTGTCTGAGATAGGCTGCGAAGAAAGACCA 1500  
QY 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCAACATACCTTGCCT 1560  
Db 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCAACATACCTTGCCT 1560  
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCCCATGATTTGCTGCTGGTTGTAAG 1620  
Db 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCCCATGATTTGCTGCTGGTTGTAAG 1620  
QY 1621 AATGGTTTACCAATCTACCTGCAAGCTACTGTCTAAATGGGGCAAGTTCTCTGATTGG 1680  
Db 1621 AATGGTTTACCAATCTACCTGCAAGCTACTGTCTAAATGGGGCAAGTTCTCTGATTGG 1680  
QY 1681 TCAGAAAACCTGGCGGAGCAAGGTCCTTAAATTTACTAGTTTCGGCACGC 1740  
Db 1681 TCAGAAAACCTGGCGGAGCAAGGTCCTTAAATTTACTAGTTTCGGCACGC 1740  
QY 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
QY 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAAAATCTCTGTTGCG 1860  
Db 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAAAATCTCTGTTGCG 1860  
QY 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGTCAT 1920  
Db 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGTCAT 1920  
QY 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTGTTCGTTGGGAACCG 1980  
Db 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTGTTCGTTGGGAACCG 1980  
QY 1981 TTCAAGAAAGACTTCAGGAGCGCTGAACTTGGACTAA 2019  
Db 1981 TTCAAGAAAGACTTCAGGAGCGCTGAACTTGGACTAA 2019

RESULT 6

AAD02799

AD AAD02799 standard; DNA; 2019 BP.

XX

AC

XX

DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.  
DE NS1; non-structure protein 1; cytotstatic; gene therapy; toxin; therapy;  
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.  
XX Parvovirus.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH 1. .2019  
FT /tag= a  
FT /product= "NS1 variant (S283A) protein"  
FT mutation replace(847, A)  
FT /tag= b  
XX EP1077260-A1.  
PN 21-FEB-2001.  
XX 13-AUG-1999; 99EP-00115161.  
PF 13-AUG-1999; 99EP-00115161.  
XX 13-AUG-1999; 99EP-00115161.  
PR (DEK-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX Nueesch J, Rommelaere J;  
PI WPI; 2001-212717/22.  
DR P-PSDB; AAY72704.  
XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.  
XX Claim 7; Page 11-14; 4lpp; English.  
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein  
CC (NS1) variant (S283A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoral diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;  
SQ Query Match 99.8%; Score 2015.8; DB 5; Length 2019;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
QY 61 AGTAAACCAAGAGTGTCTCAATTTTAAATGAAATGTTCAACTGAATGAAAA 120  
Db 61 AGTAAACCAAGAGTGTCTCAATTTTAAATGAAATGTTCAACTGAATGAAAA 120  
QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 180  
Db 121 GATATCGGATGGAATAGTTTACAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 180  
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAACGAGGACATGGAATGGGAACCACTGGAT 240  
Db 181 CGAGGAGCGGAAACTACTTTGGGACCAACGAGGACATGGAATGGGAACCACTGGAT 240  
QY 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTTAAATGTTTATTTGAA 300  
Db 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTTAAATGTTTATTTGAA 300

QY 301 GTGCTTAACAAGAAATATTTCTGGTGATGTTAAATGGTTTGTGCAACATGAATGG 360  
Db 301 GTGCTTAACAAGAAATATTTCTGGTGATGTTAAATGGTTTGTGCAACATGAATGG 360  
QY 361 GGAAGAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGACTTTAGTCAAGCT 420  
Db 361 GGAAGAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGACTTTAGTCAAGCT 420  
QY 421 CAAGGGAATTTGGTGAAGAGCAAACTAAATGTTTACTGAGAGCAGATGGTTGTTAAAGCC 480  
Db 421 CAAGGGAATTTGGTGAAGAGCAAACTAAATGTTTACTGAGAGCAGATGGTTGTTAAAGCC 480  
QY 481 TGTAAATGTCAACTAACAACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540  
Db 481 TGTAAATGTCAACTAACAACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540  
QY 541 GAGTGGGTACTCTACTTACTTATAGCATAAGCAACCAAAAAGACTATATCAAGTGT 600  
Db 541 GAGTGGGTACTCTACTTACTTATAGCATAAGCAACCAAAAAGACTATATCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATGTTCTTACTATTTTAACTHAAAAGAAAATAAGCACTAGT 660  
Db 601 GTTCTTTTGGAAACATGATGTTCTTACTATTTTAACTHAAAAGAAAATAAGCACTAGT 660  
QY 661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTACTCTGGCTGGAAACTAATCTTTTAA 720  
Db 661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTACTCTGGCTGGAAACTAATCTTTTAA 720  
QY 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780  
Db 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780  
QY 781 GTTGAACCAACAGTAACCACTGGCGAGAACTAAGCGCGAGAAATTCAAACTAAAAA 840  
Db 781 GTTGAACCAACAGTAACCACTGGCGAGAACTAAGCGCGAGAAATTCAAACTAAAAA 840  
QY 841 GAAGTTTCTATTAACACTAACAATTAAGAGCTGGTGCAATAAAGAGTAACCTCACCAGAG 900  
Db 841 GAAGTTTCTATTAACACTAACAATTAAGAGCTGGTGCAATAAAGAGTAACCTCACCAGAG 900  
QY 901 GACTGGATGATGACGAGCAGACAGTTACATTGAATGATGCTCAACACAGGTGGAGAA 960  
Db 901 GACTGGATGATGACGAGCAGACAGTTACATTGAATGATGCTCAACACAGGTGGAGAA 960  
QY 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAACCAAAACAGCA 1020  
Db 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAACCAAAACAGCA 1020  
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080  
Db 1021 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080  
QY 1081 ACAAGAGCTGCAAGATTTTGTCTTTCATGCTGGNACTATCTTAAAGTTTCCCATGCT 1140  
Db 1081 ACAAGAGCTGCAAGATTTTGTCTTTCATGCTGGNACTATCTTAAAGTTTCCCATGCT 1140  
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACTGTTTATTTTCAATGACCA 1200  
Db 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACTGTTTATTTTCAATGACCA 1200  
QY 1201 GCCAGCAGGCAAACTTATTTAGCAAGCCATAGCAACAAGCAGTTGGCAATGTGGT 1260  
Db 1201 GCCAGCAGGCAAACTTATTTAGCAAGCCATAGCAACAAGCAGTTGGCAATGTGGT 1260  
QY 1261 TGCTATTAATCGCCATGTAACCTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
Db 1261 TGCTATTAATCGCCATGTAACCTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
QY 1321 TGGTGAAGAGCTGGTAACTTTGGACAGCAGTAACCAAGTTTAAAGCAATTTGCTCT 1380  
Db 1321 TGGTGAAGAGCTGGTAACTTTGGACAGCAGTAACCAAGTTTAAAGCAATTTGCTCT 1380  
QY 1381 GGTCAAACTATTCGCTATGATCAAAAAGGAAAGGAGGAGCAACAGATTTGAACCAACCA 1440

Db 1381 GGTCAAACTATTCGCTATGATCAAAAAGGAAAGGAGGAGCAACAGATTTGAACCAACCA 1440  
QY 1441 GTCATCATGACCAACCAATGAGAACTATTACAGTGGTCAGATAGGCTGCGAAGAAAGCA 1500  
Db 1441 GTCATCATGACCAACCAATGAGAACTATTACAGTGGTCAGATAGGCTGCGAAGAAAGCA 1500  
QY 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAACTTATCATCTAACACATACCTTGCCT 1560  
Db 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAACTTATCATCTAACACATACCTTGCCT 1560  
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620  
Db 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620  
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTG 1680  
Db 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTG 1680  
QY 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTGCGCACGC 1740  
Db 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTGCGCACGC 1740  
QY 1741 TCACCAATTCACACACACCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800  
Db 1741 TCACCAATTCACACACCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800  
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATACTCTGTTGCG 1860  
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATACTCTGTTGCG 1860  
QY 1861 GGCATGCGAGAAACCCAGAACTCTGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Db 1861 GGCATGCGAGAAACCCAGAACTCTGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
QY 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980  
Db 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980  
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGCTTGAACCTTGGACTAA 2019  
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGCTTGAACCTTGGACTAA 2019  
RESULT 7  
AAT15311  
ID AAT15311 standard; DNA; 5049 BP.  
XX AAT15311;  
XX  
DT 14-OCT-1996 (first entry)  
XX  
DE Non-attenuated canine parvovirus CPV-39 passage 5 DNA.  
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;  
KW ss.  
XX Canine parvovirus.  
XX OS  
XX PN WO9614088-A1.  
XX  
PD 17-MAY-1996.  
XX  
PF 02-NOV-1995; 95WO-US014207.  
XX  
PR 08-NOV-1994; 94US-00336345.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Parrish CR, Gruenberg A, Carmichael LE;  
XX  
DR WPI; 1996-251556/25.  
XX



PT Attenuated CPV strains contg. up to 4 mutation (s) relative to control  
PT virus - useful as a veterinary vaccine against CPV disease in animals,  
XX such as wild or domestic dogs.  
PS Claim 1; Page 21-24; 42pp; English.  
XX  
CC This viral DNA is isolated from a non-attenuated CPV. The DNA is  
CC preferably derived from vB1440. The DNA is cloned into a vector which is  
CC used to transfect a host cell. The vector used is preferably pGEM3Z or  
CC pGEM5Z. The host cells to be transfected are selected from Norden  
CC laboratory feline kidney cells, mink lung cells, Madin-Darby canine  
CC kidney cells or canine A72 cells  
XX  
SQ Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;  
Query Match 51.58; Score 1040.4; DB 2; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 1.8e-269;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

1050 GAAACCGTTGAAACCCACAGTGCAGCAGCAGCAAGCAAAAGCGGGAGAAATTCAAACT 1109  
835 AAAAAAGAGTGTCTATTAATAAACTACACTTAAAGAGCTGTCATATAAGAGTAACTCA 894  
1110 AAAAGGAAGTCAATCAATCAATGTTCTTGGCGGACTTGGTTAGTAAAGAGTAACTCA 1169  
895 CCAGAGGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
1170 CTTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229  
955 GGAAGAAACCTGCTGAAATACGCTAGAGATTTGTACACTACTCTAGCAGCAACAA 1014  
1230 GGTGAAATCTTTTAAATAATACCTGAAATTTGTACTTTGCTTTAGCAGCAACAA 1289  
1015 ACAGCATTTGACTTAATTTTAAAGAAAGCTGAAACAGCAAACTAACCAACTTTTCACTG 1074  
1290 ACAGCATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1349  
1075 CTTGACACAGAGCTGACAAATTTTGTCTTTTGTGCTGGAACCTATGTTAAAGTTTC 1134  
1350 GCAATTTCTAGAACATGTCAAATTTTGTAGATGCAGGATGAAATTTGAAATTTGT 1409  
1135 CATGCTATTTGCTGTTTTTAAACAGAGGAGCAAAAGAAATTAATTAATTAATTAAT 1194  
1410 CACGCTATAGCATGTTTTTAAATAGCAAGGTGGTAAAGAAATTAATTAATTTCT 1469  
1195 GGACGAGCAGCAGCAGCAATCTATTATTGACAGCCATAGCAGCAAGCAGTTGGCAAT 1254  
1470 GACCCAGCAAGTACAGAAATCTATCTATGCTCAAGCCATAGCAGCAAGCTGGGTAT 1529  
1255 GTTGTGCTTAATGTCAGCAATGTAAATCTTTTCCATTTTAAATGCTTACCAAGAAC 1314  
1530 GTTGTGTTTAAATGTCAGCAATGTAAATTTTCCATTTTAAATGCTTACCAAGAAC 1589  
1315 TTGATTTGGTGAAGAGCTGTTAACTTTGACAGCAAGTAAACAGTTTAAAGCAAT 1374  
1590 TTAATTTGATTTGAAGAGCTGTTAACTTTGCTCAACAAATTAATCAATTTTAAAGCAAT 1649  
1375 TGCTCTGCTCAAACTATTGCTCAATGATCAAAAGAAAGGAGCAAAACAGATTGAACCA 1434  
1650 TGTCTGGAACCAATTAAGATTTGATCAAAAGGTTAAAGGAGTAAAGCAATTTGAACCA 1709  
1435 ACACAGTCTCATGACCAACAAATGAGAACATTTACAGTGGTCAAGTATGAGTGGCAAGAA 1494  
1710 ACTCCAGTAAATGACAACTAATGAAATATAACAAATTTGTGAGAAATTTGATGAGAA 1769  
1495 AGACGAGACACACTCAACCAATCAGAGCAAGATGCTTTAACTTATCATACACATACC 1554  
1770 AGACCTGAACATACACCAATTAAGAGACAGAAATTTGAAATTAAGTTAGTATGAAG 1829  
1555 TTGCTGCTGCTGCTTTGCTTGAACAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1614  
1830 CTTCCAGGAGACTTTGCTTGAACAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889  
1615 GTAAAGAAATGTTTCAATCTACATGCAAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1674  
1890 GTTAAACATGCTTTTGAATCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949  
1675 GATGCTGAGAACTGGGGGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1734  
1950 GATGCTGAGAACTGGGGGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2009  
1735 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1794  
2010 AAGACT---TAGAGACACAGGGCAAGCAATCTCAGAGTCAAGCAAGTCTTAACT 2066  
1795 CCACTTGTGATCGATCTCGAGACCTGGCTTTAGAGCTTTAGAGCTTTAGAGCAGACCAAAATCTCT 1854  
2067 CTTCTGACTCCGAGCGTAGTGGACCTTGCATCGAAGCTGGAACCGTGGAGTACTCCAGATAGCCT 2126  
1855 GTTGGGGGCACTGCGAGAAACCCAGAACACTGCGGGAAGCTGCTTCCAAAGCTGCGCAAGAT 1914  
2127 ATTGCAAGAACTGCAAAATCAACAACTTGGGCTTACTCA-----CAAGAC 2180



QY 1915 GGTCAACTGAGCCCACTTGGTCAGAGATCGAGAGGATTGAGAGGTCCTTCGGTGG 1974  
 DB 2181 GTGCAAGCCGAGTCGAGCTGCTCCGAATAGAGCGACCTGAGAGCCATCTTTACTTCT 2240  
 QY 1975 GAACCGTTGAAGAAAGACTTCACGGA 2000  
 DB 2241 GAACAATGGAAGAGATTTTCGAGA 2266

RESULT 8  
 AAT15312  
 ID AAT15312 standard; DNA; 5049 BP.  
 AC AAT15312;  
 XX  
 DT 14-OCT-1996 (first entry)  
 XX  
 DE Attenuated canine parvovirus CPV-39 passage 60 DNA.  
 XX  
 KW Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;  
 KW ss.  
 XX  
 OS Canine parvovirus.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature 59  
 FT /\*tag= a  
 FT /note= "A, C or T"  
 FT  
 FT misc\_feature 97  
 FT /\*tag= b  
 FT /note= "A, G or T"  
 FT 4745  
 FT misc\_feature 4745  
 FT /\*tag= C  
 FT /note= "A G or C"  
 FT 4881  
 FT misc\_feature 4881  
 FT /\*tag= d  
 FT /note= "A G or T"  
 FT  
 XX WO9614088-A1.  
 PN  
 XX  
 PD 17-MAY-1996.  
 XX  
 PF 02-NOV-1995; 95WO-US014207.  
 PR  
 PR 08-NOV-1994; 94US-00336345.  
 XX  
 PA (CORR ) CORNELL RES FOUND INC.  
 XX  
 XX Parrieh CR, Gruenberg A, Carmichael LE;  
 PI  
 XX WPI; 1996-251556/25.  
 DR  
 XX  
 XX Attenuated CPV strains contg. up to 4 mutation (s) relative to control  
 PT virus - useful as a veterinary vaccine against CPV disease in animals,  
 PT such as wild or domestic dogs.  
 PS  
 PS Claim 2; Page 24-27; 42pp; English.  
 XX  
 CC This viral DNA is isolated from an attenuated CPV. The DNA is preferably  
 CC derived from vB1440. The DNA is cloned into a vector which is used to  
 CC transfect a host cell. The vector used is preferably pGEM32 or pGEM52.  
 CC The host cells to be transfected are selected from Norden Laboratory  
 CC feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or  
 CC canine A72 cells  
 XX  
 SQ Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;  
 Query Match 51.5%; Score 1040.4; DB 2; Length 5049;  
 Best Local Similarity 71.4%; Pred. No. 1.8e-269;  
 Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;  
 QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60

DB 273 ATGCTCGGCAACCATGATATATCTGAGGAAGTTATGGAGGGAGTAATTTGGTTAAAGAACAT 332  
 QY 61 AGTAAACCCAGGAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGGAATA 120  
 DB 333 GCAGAAAAATGAAGCATTTTCGTTGTTTAAATGTGACAACTGCAACTAAATGGAAG 392  
 QY 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGAGCAAGCTGAAATCTTTACAA 180  
 DB 393 GATGTTTCGCTGGAACTATACCAACCAATTCAAAAATGAAGAGTAACTATCTTTAAT 452  
 QY 181 CGAGGAGCGAAACTACTTTGGGACCAAGC- - - - -GAGACATGAAATGGGAAACACA 234  
 DB 453 AGAGGAGCACAAACAGCAATGGATCAACCGAAGAAGAAATGAGCTGGGAATCGGA 512  
 QY 235 GTGGATGAAATGACCAAAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTTTA 294  
 DB 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGTGATCAATTAATAAANAATGCTT 572  
 QY 295 TTTGAAGTCTTAAACACAAAGAAATATTTCTGCTGATGTTAAATGTTGTTGTTGCAACAT 354  
 DB 573 TTTGAAGTCTTTGTTTCTAAAAATATAGAAACCAATGTTGTTGTTGTTTATTCACAT 632  
 QY 355 GAATGGGAAAGACCAAGCTGGCACTGCCATGTAATAATTGGAGGAAGGACTTTAGT 414  
 DB 633 GAATGGGAAAGATCAAGCTGGCAATGTCATGTTTCTTCATAGTAAGAACTTACAA 692  
 QY 415 CAAGCTCAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTA 474  
 DB 693 CAAGCAACTGTTAATGGCTACGACAGCAATGATATGTTGGAGTAGATGTTGGTG 752  
 QY 475 ACAGCCTGTAATGTGCAACTAACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAA 534  
 DB 753 ACTCTTTGTTGCGTAAACTTAACACCACTGAAAGATTAAGCTCAGAGAAATTCAGAA 812  
 QY 535 GACATGAGTGGTGTACTCTTACTTATAGCATAGCAACCAACCAAGAAAGCTATACC 594  
 DB 813 GATAGTAATGGGTGACTATATTAACATACAGACATAAGCAACAAAAAAGACTATGTT 872  
 QY 595 AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAAAGAAAAAAGC 654  
 DB 873 AANAATGTTTCAATTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAAAATGTC 932  
 QY 655 ACTAGTCCACCAAGAGACCGAGCTATTTTTCAGCAGTACTCTCGCTGGAATACTAAC 714  
 DB 933 CATAT- - -GACAAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGTTGGAATTTAAC 989  
 QY 715 TTTTAAAGAGGCGGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCA 774  
 DB 990 TTTATGAAGTATCAAGACAGACAAATTTGCAGCACACTTTTACACTGAACAAATGAACCA 1049  
 QY 775 GAAACGGTTGAAACCAACAGTAACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACT 834  
 DB 1050 GAAACGGTTGAAACCAACAGTAACTGCGCAGGAAACTAAGCGCGGAGAAATTCAAACT 1109  
 QY 835 AAAAAAGAGTGTCTTATTTAAACTACACTTAAAGAGCTGGTGATAAAAAGAGTAACCTCA 894  
 DB 1110 AAAAAAGAGTGTCTTATTTAACTCAATCAATGTTCTTTGCGGACTTGGTTAGTAAAGAGTAACTCA 1169  
 QY 895 CCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
 DB 1170 CCTGAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229  
 QY 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCAGAACCAAA 1014  
 DB 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTAGCAGAACAAA 1289  
 QY 1015 ACAGCATTTGACTTAAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTG 1074  
 DB 1290 ACAGCATTTGAAATTAATCTTGAAGAAAGCAGATAATTAATACTAACTAACTTTGATCTT 1349  
 QY 1075 CTTGACACAAAGAGCTGACAGAAATTTTGTGCTTTTCATGGCTGGAATATGTTAAAGTTGC 1134

```
Db 1350 GCAAAATTCAGACATGTCAAATTTTGAATGTCACGGATGGAATGGATTAAAGTTGT 1409
Qy 1135 CATGCTATTGCTGCTTTTAAACACAGAGAGGCAAAAGAAATACTGTTTATTTCAT 1194
Db 1410 CAGCTATTAGCATGTGTTTAAATAGACAGGTTGTTAAAGAAATACAGTCTCTTTTCAT 1469
Qy 1195 GGACCAAGCAGCAGCAGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGSCAAT 1254
Db 1470 GGNACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1529
Qy 1255 GTTGGTTGCTATATGATGACGACCAATGAACTTTCCATTTTAAATGATGCTGACCAACAGAAC 1314
Db 1530 GTTGGTTGTTAATGATGACGACCAATGAAATTTTCCATTTTAAATGATGCTGACCAATAAAT 1589
Qy 1315 TTGATTTGGGTAGAGAGAGCTGTTACTTTTGACAGCAAGTAAACAGTTTAAAGCAAT 1374
Db 1590 TTAATTTGGATTGAAGAGCTGTTACTTTTGACAGCAAGTAAACAGTTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTGCTGATGATCAAAAGGAAAGGAGCAGCAAAACAGATTGAACCA 1434
Db 1650 TGTCTTGGACAACTATTGAATTTGATCAAAAGGTAAGGAAATTTGAACCA 1709
Qy 1435 ACACGATCATATGACCAATGACAAATGACATTTACAGTGTGATGATGCTGCGAGAA 1494
Db 1710 ACTCCAGTAATATGACAACTAATGAAATATTAACAAATTTGAGAAATTTGATGTGAAGA 1769
Qy 1495 AGACCAAGACACTCAACCAATGACAGACAGAAATGCTTAAACATTCATCTAACACATACC 1554
Db 1770 AGACCTGAACATACACCAATTAAGACAGAAATTTGACAAATTTAGTATGTTAG 1829
Qy 1555 TTGCTCTGGTACATTTGTTGTTGTTGACAAATTTGAATGCGCCCATGATTTGCTGCTG 1614
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGATTAAGAAAGATGGCTTTAATATGTCATGTTA 1889
Qy 1615 GTAAAGATGTTACCAATCTACATGCGCAAGTACTGTGCTAAATGGGGCAAGTTCTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACACATCATTTGGGGAAAGTACCA 1949
Qy 1675 GATTGTCAGAAACTGGGGGAGCCAAAGTGCCCAACTCTATAAATTTACTAGTTCTG 1734
Db 1950 GAATGGATGAAACTGGGGGAGCCCTAAATACAGAGGATATAATTCACAGGTTGC 2009
Qy 1735 GCACGCTCACATTCAGCACACGCAAAAGTACCGCTCTCAGCCAGAACTATGCACTAAT 1794
Db 2010 AAAGACT--TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAAT 2066
Qy 1795 CCACTTGCATCGATCTCAGGACCTGGCTTTAGAGCTTGGAGCACACCAATATCTCT 1854
Db 2067 CCTCTGACTCCGACGTTAGTGGACCTTGCACCTGGAACCGTGGAGTACTCCAGATACGCCT 2126
Qy 1855 GTTGGGGCACTGCAGAAACCCAGAACACTGGGGAGAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAAACTGCAATCAACATCAACATCAACCAACTTGGCTTACTCA-----CAAGAC 2180
Qy 1915 GGTCACTGAGCCCACTCTGCTCAGAGATCAGAGAGATTTAGAGCGTCTTCGGTGGG 1974
Db 2181 GTCAAGCGAGTCCGACGTTGTCGAAATAGAGCGACCTGAGAGCCATCTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAAGCTTCAGGA 2000
Db 2241 GAACAATTTGAAGAAGATTTTCGAGA 2266
```

## RESULT 9

AAT88324 standard; DNA; 5049 BP.

AC AAT88324;

XX AC

DT 17-OCT-2003 (revised)

DT 21-MAY-1998 (first entry)

XX DE Attenuated canine parvovirus (vBI440) genomic DNA.

XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;  
KW feline panleukopenia virus; mink enteritis virus; infection; ds.  
XX Canine parvovirus; vBI440 (ATCC VR 2489).

XX Key Location/Qualifiers  
FH mutation 59  
FT /tag= c  
FT /note= "base 59 is G in CPV-39 (passage 5)"  
FT mutation 97  
FT /tag= d  
FT /note= "base 97 is C or T in CPV-39 (passage 5)"  
FT CDS 273..2279  
FT /tag= a  
FT /note= "NS1/NS2 coding region"  
FT CDS 2286..4541  
FT /tag= b  
FT /note= "VP1/VP2 coding region"  
FT mutation 4745  
FT /tag= e  
FT /note= "base 4745 is T in CPV-39 (passage 5)"  
FT mutation 4881  
FT /tag= f  
FT /note= "base 4881 is C in CPV-39 (passage 5)"

XX WO9742972-A1.

XX 20-NOV-1997.

XX 06-MAY-1997; 97WO-US007584.

XX 15-MAY-1996; 96US-00647655.

XX (CORR ) CORNELL RES FOUND INC.

XX Parrish CR, Carmichael LE, Gruenberg A;

XX WPI; 1998-008583/01.

XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as  
vaccines for protection against parvovirus and feline pan-leukopenia  
virus infections.

XX Example 8; Page 34-37; 60pp; English.

XX This DNA sequence comprises an attenuated virus genome derived by serial  
passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate  
39 in NUPK feline kidney host cells. The attenuated virus is designated  
vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence  
(see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations  
are within the hairpin formed by the 3' terminal palindromic region: the mutation  
at nucleotide 59 introduces an A into a G-C rich region within the tip of  
the hairpin, disrupting the base pairing in one of the 2 small internal  
palindromes within that sequence; the thymine at nucleotide 97 is  
adjacent to the mismatched bubble (flip-flop) sequence within the  
palindrome. The DNA from attenuated CPV strains (see also AAT88321) is  
used for the production of infectious molecular DNA clones, which, in  
turn, can be transfected into cells to generate master stocks of the  
virus. The attenuated viruses can be used in dogs as a vaccine to protect  
against CPV disease, or more generally in cats and minks to protect  
against feline panleukopenia virus and mink enteritis virus. The vaccines  
protect against the currently prevalent CPV-2b type (and all extant  
strains of types 2 and 2a), providing a long term immune response.  
(Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

XX Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

XX Best Local Similarity 71.4%; Pred. No. 1.8e-269;

XX Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAGGAAAA 60

Db 273 ATGCTGGCAACCAAGTATCTGAGGAAGTTATGGAGGAGTAATTTGGTTAAAGAACAT 332  
Qy 61 AGTAACAGGAAGTGTCTCATTTGTTTAAATAAGAAATGTTCAACTGAATGGAAAA 120  
Db 333 GCAGAAAAATGAAGCAATTTGTTGTTTAAATGTCAGAACGTCACACTAAATGGAAAG 392  
Qy 121 GATATCGGATGGATAGTTACAAAAAGAGCTCGAGGAGACGAGCTGAATCTTTTACAA 180  
Db 393 GATGTTGCGTGGAAACAACTATACCAACCAATTCAAAATGAAGAGCTAACATCTTTAAT 452  
Qy 181 CGAGGAGCGGAATCTCTCGGACCAAGC-----GAGGACATGGAATGGGAAACACCA 234  
Db 453 AGAGGAGCACAACAGCAATGGATCAACCCGAAGAGAAATGAGCTGGGAATCGAA 512  
Qy 235 GTGGATGAATGACCAAAAAAGCAAGTATTCATTTTGTGTTGTTTAAATAAGTTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAAACTTTTGTGATCAATTAATTAATAAATGTC 572  
Qy 295 TTTGAAGTCTTAAACAAAGAAATATATTTCTGGTGGTATGTTAATTTGGTTTGTGCAACAT 354  
Db 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAAAATGAATGTTGTTTGTGTTTATTCACAT 632  
Qy 355 GAATGGGAAAGACCAAGCTGCGCATGTACTAATTCGAGGAAAGCACTTTAGT 414  
Db 633 GAATGGGAAAGATCAAGCTGGCATGTGATGTTTCTTCATAGTAAGAACTTACAA 692  
Qy 415 CAAGCTCAAGGGAATGTTGGAGGAAGCAACTAAATGTTTACTCGAGCAGATGTTGGTA 474  
Db 693 CAAGCAACTGTTAATGGCTACGACAGCAAAATGATATGTTTGGAGTAGATGTTGGTG 752  
Qy 475 ACAGCTGTAAATGTGCACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534  
Db 753 ACTCTTTGTTGCGTAAACTTAAACCAACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812  
Qy 535 GACAATGAGTGGTACTCTACTTACTTATAAGCATAGCAACCAACCAAGCAAGTACTTACC 594  
Db 813 GATAGTAATGGGTGACTATTAACATACAGACATAAGCAAAACAAAGAAAGTACTATGTT 872  
Qy 595 AAGTGTGTTCTTTTGGAAACATGATTCCTTACTATTTTAACTAAAAAGAAATAAGC 654  
Db 873 AATAGTTCATTTTGGAAATATGATGATATATCTTTTAAACAAAGAAATAATGTC 932  
Qy 655 ACTAGTCCACCAAGACGAGGCTATTTCTTGTAGAGTGAATCTGCTGGGAAAACTAAC 714  
Db 933 CACAT---GACAAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGGTGGAAATTAAC 989  
Qy 715 TTTTAAAGAGCGGCGCCATCTAGTCAGCAAACTATACACTGATGACATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAGCACACTTTTACACTGAACAAATGAAACCA 1049  
Qy 775 GAAACGGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGGAGAAATTCAAACT 834  
Db 1050 GAAACGGTTGAAACCAAGTAAACACTGCGCAGCAGCAAGCAAGCGCGGAGAAATTCAACT 1109  
Qy 835 AAAAAAGAAAGTTCTATTAATAACTACATTTAAAGAGCTGGTGCATTAAGAGAGTAACCTCA 894  
Db 1110 AAAAAAGAAAGTCAATCAATGATCTTTGCGGCACTTGGTTAGTAAAGAGTAACATCA 1169  
Qy 895 CCAGAGGACTGGATGATGATGAGCCAGCAGACAGTTTACATTTGAATGATGGCTCAACAGGT 954  
Db 1170 CTTGAAGACTGGATGATGTTTACAAACAGATAGTTATTTGAAATGATGGCAACACAGGA 1229  
Qy 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAA 1014  
Db 1230 GGTGAAATCTTTTAAAAAATACCTTGAATTTGTACTTTGACTTTTAGCAAGAACAAA 1289  
Qy 1015 ACAGCATTTGACTTAATTTTAAAGAAAGCTGAAACCAAGCAAGCAAACTAACCCTTTTCACTG 1074  
Db 1290 ACAGCATTTGAATTAATCTTGAAGAGCAGATTAATACTAACTAACTAACTTTGATCTT 1349  
Qy 1075 CTTGACACAAAGCCTGCAGAAATTTTGTCTTTTATGCTGGAACTATGTTAAAGTTTGC 1134

Db 1350 GCAAATTTAGACATGTCAAATTTTATAGATGCCGAGTGAATTTGGATTTAAAGTTGT 1409  
Qy 1135 CATGCTATTTGCTGTGTTTTTAAACACAGAGGAGCAAAAGAAATACTGTTTTTATTTTCAT 1194  
Db 1410 CAGCTATAGCATGTGTTTTTAAATAGCAAGGTGGTAAAAAGAAATACAGTTCTTTTTTCAT 1469  
Qy 1195 GGACCAAGCAGCAGCAGCAAAATCTATTTATGCAAGCCATAGCAGAGTTGGCAAT 1254  
Db 1470 GGACCAAGCAGTACAGGAAATCTATCTGCTCAAGCCATAGCAAGCTGTGGGTAAT 1529  
Qy 1255 GTTGGTGTCTATAATCAGCCCAATGTAAACTTTTCCATTTTAACTGTGTACCAACAAGAAC 1314  
Db 1530 GTTGGTGTATAATGCAAGCAATGTAAATTTTCCATTTTAACTGTGTACCAATAAAT 1589  
Qy 1315 TTGATTTGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCAAT 1374  
Db 1590 TTAATTTGGATGGAAGAGCTGGTAACTTTGGTCAACAAAGTTAATCAATTTAAAGCAATC 1649  
Qy 1375 TGCTCTGCTCAAACTATTGCGCATTTGATCAAAAGGAAAGGCAAGCAAGATTGAACCA 1434  
Db 1650 TGTCTGGAACAACAATTAAGATTGATCAAAAGGTAAGGAAAGTAAAGCAAAATTGAACCA 1709  
Qy 1435 ACACCACTCATCATGACCAACAATGAGAACATTACAGTGTGTCAGAACTAGCTCGGAAGAA 1494  
Db 1710 ACTCCAGTAAATATGACAACTAATGAATAATTAACAATTTGTGAGAAATTTGATGTGAAG 1769  
Qy 1495 AGACCAAGAACACACTCAACCAATTCAGACAGCAAGTCTTAAACATTCATCTAACACATACC 1554  
Db 1770 AGACCTGAACATACACAAACCAATTAAGAGACAGAAATGTTGAACATTAAGTTAGTATGAAG 1829  
Qy 1555 TTGCTCTGCTGACTTTGTTGTTTGGTTGACAAAATGAATGSCCCATGATTTGTGCTGGTTG 1614  
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGTTGTAAGAGAAATGSCCTTTAATATATGTGATGGTTA 1889  
Qy 1615 GTAAAGAAATGTTTACCAATCTACATGSCAAAGCTACTGTGCTAAATGGGCAAAAGTTCTCT 1674  
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTACTATACATCATTTGGGGAAGAAAGTACCA 1949  
Qy 1675 GATGCTCAGAAAACTGGCGGAGCCAAAGGTGCGCAACTCTTATAAATTTACTAGGTTG 1734  
Db 1950 GAATGGATGAAAACTGGCGGAGCTTAAATACAAAGAGGTATAAATTCACCAGGTTGC 2009  
Qy 1735 GCAGCTCAGCAATTCAGACACCGAAAGTACGCTCTCAGCCAGCAACTATGCACTAACT 1794  
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT 2066  
Qy 1795 CCACCTTTCATCGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAAAATCTCTCT 1854  
Db 2067 CTTCTGACTCCGACGCTAGTGGACCTTGACCTGGAACCGTGGAGTACTCCAGATAGCCT 2126  
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGAACACTTGGGGAAGCTGGTTCCAAAGCTTCCCAAGAT 1914  
Db 2127 ATTGCAAGAACTGCAATCAACATCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180  
Qy 1915 GGTCAACTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGGCTGCTTCGGTGGC 1974  
Db 2181 GTGCAAGCAGTCCGACGCTGGTCCGAAATAGAGGCAAGCTCAGAGCCATCTTTACTTCT 2240  
Qy 1975 GAACCGTTGAAGAAAGACTTCAGCGA 2000  
Db 2241 GAACAATTTGAAGAAAGATTTTCGAGA 2266

RESULT 10  
AAT88321  
ID AAT88321 standard; DNA; 5049 BP.  
XX  
AAT88321;  
XX  
21-MAY-1998 (first entry)  
XX  
Attenuated canine parvovirus genomic DNA.  
DE  
XX





393	Db	 GATGTTTCGCTGGAACTACTCTGGGACCAAGC-----GAGGACATGGAATGGAAGAGCTAAATCATCTTTAAATT	452
181	Qy	CGAGGAGCGAACTACTCTGGGACCAAGC-----GAGGACATGGAATGGAAGAGCTAAATCATCTTTAAATT	234
453	Db	AGAGGAGCACAACAGCAATGGATCAACCGAAGAGAGAAATGGACTGGGAATCGGAA	512
235	Qy	GTGGATGAAATGACCAAAAAGCAAGTATTCATTTTTTGATTTCTTTGGTTAAAAAATGTTTA	294
513	Db	GTTGATAGTCTCGCAAAAAGCAAGTACAACTTTTGATGTCATTAATTAAGAAATGTCIT	572
295	Qy	TTTGAAGTCTTAACACAAGAAATATATTTCTCTGGTGATGTTAATTTGGTTTGGCAACAT	354
573	Db	TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAAAATGAATGTGTGTTGTTTATTTCAACAT	632
355	Qy	GAATGGGAAAAAGACCAAGCTGGCACTGCCATGTACTAATTCGGAGGAAGGACTTTAGT	414
633	Db	GAATGGGAAAAAGATCAAGGCTGGCATTTGTCATTTTACTTCATAGTAAGAATCTACAA	692
415	Qy	CAAGCTCAAGGGAAATGGTGAGAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTA	474
693	Db	CAAGCAACTGGTAAATGGCTACGAGACAAATGAAATATGTATTTGGAGTAGATGTTGGTG	752
475	Qy	ACAGCCTGTTAATGTGCAACTTAAACCAAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAA	534
753	Db	ACTCTTTGTTTCGGTAAACTTTAAACCAACTGAAAGAAATTAAGCTTCAGAGAAATTCAGAA	812
535	Qy	GACAAATGAGTGGTTCCTACTCTACTTACTTATTAAGCATTAAGCAAAACCAAAAGAACTATACC	594
813	Db	GATAGTGAATGGGTGACTATATTAAACATACAGACATAAGCAAAACCAAAAGAACTATGTT	872
595	Qy	AAGTGTCTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAAGC	654
873	Db	AAAATGGTTCATTTTGGAAATATGATAGCATATTACTTTTAAACAAGAAATAAATGTC	932
655	Qy	ACTAGTCCACAAGACGAGGCTATTTTCTTTAGCAGTGACTCTCGCTGGAAAACTAAC	714
933	Db	CACAT--GACAAAAGAAAGTGGCTATTTTTTAAGTACTGATTTCTGTTGGAATTTAAC	989
715	Qy	TTTTTAAAGAGCGGAGCGCCACTAGTACAGCAAACTATACACTGATGACATCGCGCCA	774
990	Db	TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGAAACCA	1049
775	Qy	GAAACGGTTGAAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT	834
1050	Db	GAAACCGTTGAAACCAACAGTGAACGACAGACAGGAAACAAAGCGCGGAGAAATCAAAC	1109
835	Qy	AAAAAGAGGTTCTATTAAAACTACACTTTAAAGAGCTGGTGCAATAAGAGGTAACTCTCA	894
1110	Db	AAAAAGAGGTTGATCAATCAAAATGTACTTTTGGGAGCTTGGTTAGTAAAGAGGTAAACATCA	1169
895	Qy	CCAGAGACTGGATGATGATGACGACGACAGTATCAATGCAAAATGATGGCTCAACCAAGT	954
1170	Db	CCTGAAGACTGGATGATGTTTAAACCAAGATAGTTATATTGAATGATGGCACAACCAAGGA	1229
955	Qy	GGAGAAAACCTGTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCGAGAACCAAA	1014
1230	Db	GGTGAAAAATCTTTTAAAAAATACTTTGAAATTTGTACTTTTGACTTTAGCAAGAACAAA	1289
1015	Qy	ACAGCAATTTGACTTTAATTTTAAAGAAAAAGCTGAAACCAAGCAAACTTAACCACTTTTCACTG	1074
1290	Db	ACAGCAATTTGAAATTAATACTTGAAGAAAGCAGATAATACTTAACTAACTAATCTTTGATCTT	1349
1075	Qy	CCTGCACAAGAGCCTGCAGAAATTTTTCCTTTTCACTGGCTGGACATATGTTAAAGTTTGC	1134
1350	Db	GCAAAATCTAGAACATGTCAAATTTTTTAGAATGCAAGGATGGAAATTTGGATTTAAAGTTTGT	1409
1135	Qy	CATGCTATTTGCTGTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATCTGTTTTATTTCAT	1194
1410	Db	CACGCTATAGCATGTGTTTTTAAATAGACAGGTGTTAAAGAAATACAGTTCCTTTTCAT	1469
1195	Qy	GGACCAAGCCAGCAGCGAAATCTATTATTGCAACAGCCATAGCACAAAGCAGTTGGCAAT	1254

Db	1470	GGACCAAGTACAGGAAATCTATCTATGCTCAAGCCATAGCAACAAGCTGTGGTAAT	1529
Qy	1255	GTTCGTTGCTATAATGCAGCAATGTAAACTTTCCATTTAAATGACCTGTACCAACAAGAAC	1314
Db	1530	GTTCGTTGTTATATGCAGCAATGTAAATTTCCATTTAAATGACCTGTACCAATAAAAT	1589
Qy	1315	TTGATTTGGGTAGAGAAGCTGGTAACTTTTGGGACAGAAAGTAAACAGTTTAAAGCAAT	1374
Db	1590	TTAATTTGGATTTGAAGAAGCTGGTAACTTTGGTCAACAGTTAATCAATTTAAAGCAATC	1649
Qy	1375	TGCTCTGGTCAAACTATTCCATTTGATCAAAAAGGAAGGACGAGAAACAGATTTGAACCA	1434
Db	1650	TGTTCTGGACAAACAAATTTAGATTTGATCAAAAAGGTAAAGGAAGTAAAGCAATTTGAACCA	1709
Qy	1435	ACACGCTCATCATGACCAACAAATGAGACATTACAGTGGTCAAGATAGGCTCGCAAGAA	1494
Db	1710	ACTCCAGTAATTTAGCACTAATGAAATTTAAACAATTTGTGAAATTTGATGTGAAGAA	1769
Qy	1495	AGACCAAGACACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTCATCTAACACATACC	1554
Db	1770	AGACCTGAACATACACAAACCAATAAGAGACAGAAATGTTGAACATTAAAGTTAGTATGTAAG	1829
Qy	1555	TTGCTCTGGTGAATTTGGTTTGTGACAAAATGAATGGCCCAATGATTTGTGCTTGGTTG	1614
Db	1830	CTTCCAGGAGACTTTGGTTTGGTTGTAATAAGAAAGAAATGGCCCTTTAATATGTGCATGGTTA	1889
Qy	1615	GTAAGAATGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAGTTCCCT	1674
Db	1890	GTTAAACATGTTTTGAAATCAACATGGCTAACTATACATCATTTGGGGAAAGTACCA	1949
Qy	1675	GATTGGTCAGAAAATCTGGGGCGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGGTTGC	1734
Db	1950	GAATGGGATGAAACCTGGGGCGAGCCCTAAATAACAAGAGGTATAAATTCACCGGTTGC	2009
Qy	1735	GCAGCTCACCATTCAGGACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCATAACT	1794
Db	2010	AAAGACT---TAGAGACACAAAGCGGCAAGCAATCCTCAGAGTCAGAACCAAGTTCTAACT	2066
Qy	1795	CCACTTCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGACACACCAATACTCCT	1854
Db	2067	CCTCTGACTCGGACGTAGTGGACCTTTGCACTGGAACCGTGGAGTACTCCAGATAGCCT	2126
Qy	1855	GTTCGGGGCACTGCAGAAACCCAGAACACTTGGGGAAAGCTGGTTCCAAAGCCTGCCAAGAT	1914
Db	2127	ATTGCAAACTGCAATCAACAATCAACCACTTGGCGTTACTCA-----CAAAGAC	2180
Qy	1915	GGTCAACTGAGCCCACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTGTCTTCGGTGGC	1974
Db	2181	GTGCAAGCGAGTCCGACGTGGTCCGAATAGAGCGAGACCTGAGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTTGAAGAAAGACTTCAGCGA	2000
Db	2241	GAACAATTTGGAAGAAGATTTTCGAGA	2266
RESULT 12			
AAN40252 standard; DNA; 3524 BP.			
ID	AAN40252 standard; DNA; 3524 BP.		
XX	AAN40252;		
XX	AC		
DT	24-OCT-2003 (revised)		
DT	12-JAN-1992 (first entry)		
XX	Sequence from the double-stranded replicative form DNA of porcine parvovirus.		
DE	Protein envelope; immunogen; vaccine; antigen; epitope; ds.		
XX	Porcine parvovirus; NADL-2 virulent strain.		
OS	Porcine parvovirus; NADL-2 virulent strain.		
XX	Key		
XX	FH		
FT	CDs		
XX	Location/Qualifiers		
XX	1. .2073		

```
FT      /*tag= a  
FT      /note= "see AAP40306"  
FT CDS 2107..3522  
FT      /*tag= b  
FT      /note= "see AAP40675"  
XX  
PN W08402847-A.  
XX  
XX 02-AUG-1984.  
XX  
XX 19-JAN-1984; 84WO-US000063.  
XX  
XX 19-JAN-1983; 83US-00459203.  
PR 06-JAN-1984; 84US-00567968.  
XX  
XX (AMGE-) AMGEN.  
XX PA  
XX XX  
XX PI  
XX XX  
XX XX  
DR WPI; 1984-201354/32.  
DR P-PSDB; AAP40306, AAP40675.  
XX  
XX Polypeptide obtd. by recombinant DNA methods - for vaccination against  
XX parvovirus infections in man and animals.  
XX  
XX Claim 10; Table II, Page 33-49; 80pp; English.  
XX  
XX The inventors claim an immunologically active polypeptide for the  
XX development of vaccinal immunity against parvovirus infection. Also  
XX claimed are DNA sequences wholly or partly duplicative of defined  
XX sequences. The polypeptides are used in vaccines for conferring  
XX protection against parvovirus infections in man and animals. (Updated on  
XX 24-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;  
  
Query Match 44.6%; Score 901; DB 1; Length 3524;  
Best Local Similarity 69.7%; Pred. No. 6e-232;  
Matches 1296; Conservative 0; Mismatches 545; Indels 18; Gaps 5;  
  
QY 4 GCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAAAGT 63  
DB 13 GCGGAAACACTTACTCGGAAGAGGTACTAAAGCTTACCACTGGCTTCAAGATATGCT 72  
QY 64 AACCAGGAAGTGTCTCATTTGTTTTTAAATAATGAAATGTTCAACTGAATGGAAGAAT 123  
DB 73 CAAAAGAAGCAATCTCTTATGATTTTAAACACAAAAGTCAATCTAATGGAAGAA 132  
QY 124 ATCGGATGGAATAGTTACAAAAGAGCTGCAGAGAGCAGCTGAAATCTTTACACGA 183  
DB 133 ATTGCTTGAATAACTACAAAGATACACAGATGCGGAAATGATAAACCTTACAAAG 192  
QY 184 GGAGCGAATCTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGATGAA 243  
DB 193 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCGACAGC 252  
QY 244 ATGACCAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAAGTG 303  
DB 253 CTCACAAAAGCGCAAGTACTGATTTTGTGCTCTCTTTTAAATAATGCTCTTTGAAGGT 312  
QY 304 CTTAAACAAGAATAATATTTCTGGTGATGTTAATTTGTTTGTGCAACATGAATGGGA 363  
DB 313 ATATTGCAAAAGAACCTTAAGTCCAAGTGACTGCTTCTTACAGCATGAACATGGT 372  
QY 364 AAGACCAAGCTGGCACTGCCATGTTACTAATTTGGAGGAAGGACCTTTAGTCAAGTCAA 423  
DB 373 CAAGATACCTGGCTATCACTGCCATGTACTAGTGGGAAAGGCTTTACAAAGCAATG 432  
QY 424 GGGAAATGGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCTGT 483  
DB 433 GGAATATGGTTACGAAGAACAATTAACAATTTATGGATGATGGTTGAATATCAATGC 492  
QY 484 AATGTGCAACTAACACCAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAAGCAATGAG 543
```

```
Db 493 AAAGTACCTCTAACACCAGTTGAAAGAAATAAAATTAAGGGAATTAGCAGAGGATGGTGAG 552  
QY 544 TGGGTTACTCTTACTTATAGCAATAAGCAAAACCAAAAAAGACTATATACCAAGTGTGTT 603  
Db 553 TGGGTATCGTACTACTAACCTACACTCACAAACAACTAAAAAACAATATACAAAAATGACT 612  
QY 604 CTTTTTGGAAAAATGATGCTTACTATTTTAACTAAAAAGAAAAATAGCAGCTAGTCCA 663  
Db 613 CATTTTGGAAATATGATGCTTACTACTCTCTAAATAAAAAAAGAAGCAACT----- 666  
QY 664 CCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGGAAAACTAACTTTTAAAA 723  
Db 667 GAAAGAGAGCATGGATATATCTCAGCTCAGATTTCTGGCTTCTATGACAAATTTCTTAAA 726  
QY 724 GAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCAGAAAAACGGTT 783  
Db 727 GAAGCGAGAGACACTTAGTCAGTCACCTATTTTACTGAAGCAATAAACCCTGAACCTGTG 786  
QY 784 GAAACACAGTAAACCACTGCGGAGGAAACTAAGCGGCGGAGAAATTCAAACTAAAAAGAA 843  
Db 787 GAAACCAACGGTTACTACAGCTCAGGAAGTTTCCCGAGGCGAGAAATACAAAACAAAAAGAA 846  
QY 844 GTTCTCTTAAACTACACTTAAAGAGCTGCTGCATATAAAGAGCTAACCTCACCAGAGGAC 903  
Db 847 GTAAGCATAAATGCAATAAGAGACTTGGTTAATAAAGATGTACTAGCATAGAAGGC 906  
QY 904 TGGATGATGTCAGCGACAGCTTACATTTGAAATGATGCTCAACACAGGTGGAGAAAAAC 963  
Db 907 TGGAGTATGACAGATCCAGACAGTTATATAGAAATGATGGCTCAACCGGAGGAGAAAT 966  
QY 964 CTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCATTT 1023  
Db 967 TTAATCAAAAATACACTAGAAATAACAACTCTTACTCTAGCAAGAACAAAAACAGCATAT 1026  
QY 1024 GACTTAAATTTAGAAAAAGCTGAAACCAAGCAACTAACCAACTTTTCACCTGCCTGACACA 1083  
Db 1027 GACTTAAATCTTGAAGGCAAAACCAAGCATGCTACCAACATTTAATAATATTAGCAATACA 1086  
QY 1084 AGAGCTGTCAGAAATTTTGTCTTTCATGCTGGAACCTATGTTAAAGTTTGCCTATGCTATT 1143  
Db 1087 AGAACATGTAAATATTTTCAGCATGCAATTTGGAACCTACATTAAAGTCTGCCATGCTATA 1146  
QY 1144 TGCCTGTGT-TTTAAACAGACAAGGAGGCAAAAGAAATA-----CTGTTTATTTTCATGGA 1197  
Db 1147 ACTTGTGTACTGAAACAGACAAGGAGGAAAAAGAAATAACAATTTCTATTTTCATGTCATGGG 1206  
QY 1198 CAGCCAGCACAGGCAAAATCTATTATTGACAGCCATAGCACAGCAGTGGCAATGTT 1257  
Db 1207 CCAGCATCAACAGGAAAAAGATTAATTTGCTCAACACATTTGCAAACTTAGTTGTAATGTT 1266  
QY 1258 GGTGTGCTATAATGACGCAATGTAAACTTTTCCATTTTAACTGCTGTACCAACCAAGAACTTG 1317  
Db 1267 GGTGTGCTACATGACGCAATGTGAATTTCCATTTTAACTGCTGTACCAATTAANAACCTTA 1326  
QY 1318 ATTTGGGTGAGAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGC 1377  
Db 1327 ATATGATTTGAAGAAGCAGGAAACTTCTCTAACCAAGTAAACCAATTTCAAAGCCATATGT 1386  
QY 1378 TCTGTGCAAACTTATTCGCATTGATCAAAAAGGAAAGGCGGCAAAACAGATTGAACCAACA 1437  
Db 1387 TCAGGTCAAACTATTAAGAAATTCACAAAAGGTAAAGGAGGCAAAACANAATTTGAACCAACT 1446  
QY 1438 CCAGTCATCATGACCAAAATGAGAACATTTACAGTGTGTAGAAATAGCTGCGGAAAGAAAGA 1497  
Db 1447 CTTGTAATTAATGACTACAAATGAAGACATACTTAAGTTTGAATAGGATGCGGAGGAAAGA 1506  
QY 1498 CCAGAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCAATCTAAACACATCCTTG 1557  
Db 1507 CCAGAACATACACAACCAATAGAGACAGAAATGTTAAACATACACCTTAACCAAGAAACTG 1566  
QY 1558 CTTGTGTGACTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTTA 1617
```









GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds  
(without alignments)  
8723.907 Million cell updates/sec

Title: US-10-069-056-8  
Perfect score: 2019  
Sequence: 1 agcgtggaatgcttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/5A-COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B-COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A-COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B-COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS-COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1 Sequence 1, Appli
2	1040.4	51.5	5049	2	US-08-336-345-2 Sequence 2, Appli
3	1040.4	51.5	5049	1	US-08-647-655-1 Sequence 1, Appli
4	1040.4	51.5	5049	2	US-08-647-655-2 Sequence 2, Appli
5	109.4	5.4	4680	1	US-08-254-358-1 Sequence 1, Appli
6	109.4	5.4	4680	1	US-08-475-391-1 Sequence 1, Appli
7	109.4	5.4	4680	2	US-08-709-609-1 Sequence 1, Appli
8	109.4	5.4	4680	5	PCT-US95-07178-1 Sequence 1, Appli
9	109.4	5.4	4681	4	US-09-807-802A-18 Sequence 18, Appli
10	109.4	5.4	4910	2	US-08-331-384-2 Sequence 2, Appli
11	109.4	5.4	4910	2	US-08-836-087-2 Sequence 2, Appli
12	109.4	5.4	4910	3	US-09-246-320-2 Sequence 2, Appli
13	109.4	5.4	4910	3	US-09-546-738-2 Sequence 2, Appli
14	109.4	5.4	7214	4	US-09-438-268-1 Sequence 1, Appli
15	109.4	5.4	7557	4	US-09-770-315-3 Sequence 3, Appli
16	109.4	5.4	8151	4	US-09-438-268-2 Sequence 2, Appli
17	109.4	5.4	8179	4	US-09-438-268-5 Sequence 5, Appli
18	109.4	5.4	8698	4	US-09-770-315-2 Sequence 2, Appli
19	104.6	5.2	939	4	US-09-532-594B-12 Sequence 12, Appli
20	104.6	5.2	1197	4	US-09-532-594B-13 Sequence 13, Appli
21	104.6	5.2	1611	4	US-09-532-594B-14 Sequence 14, Appli
22	104.6	5.2	1872	4	US-09-532-594B-3 Sequence 3, Appli
23	104.6	5.2	1872	4	US-09-532-594B-15 Sequence 15, Appli
24	104.6	5.2	4767	4	US-09-532-594B-1 Sequence 1, Appli
25	99.8	4.9	969	4	US-09-807-802A-10 Sequence 10, Appli
26	99.8	4.9	1200	4	US-09-807-802A-8 Sequence 8, Appli
27	99.8	4.9	1641	4	US-09-807-802A-6 Sequence 6, Appli

28	99.8	4.9	1872	4	US-09-807-802A-4	Sequence 4, Appli
29	99.8	4.9	4883	4	US-09-807-802A-19	Sequence 19, Appli
30	99.8	4.9	4718	4	US-09-807-802A-1	Sequence 1, Appli
C 31	64.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appli
C 32	61.2	3.0	1141	4	US-09-806-708B-22	Sequence 22, Appli
C 33	44	2.2	1141	4	US-09-806-708B-22	Sequence 22, Appli
C 34	40.8	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 35	40.8	2.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
36	40.2	2.0	277	3	US-09-007-005-3	Sequence 3, Appli
37	40.2	2.0	277	3	US-09-244-796-3	Sequence 3, Appli
38	38.6	1.9	116592	4	US-09-818-512-3	Sequence 2813, Ap
C 39	38.4	1.9	832	4	US-09-621-976-2813	Sequence 14265, A
40	38.4	1.9	1367	4	US-09-270-767-14265	Sequence 1, Appli
41	38.4	1.9	6328	3	US-08-913-832A-1	Sequence 1, Appli
42	38.4	1.9	6328	4	US-09-249-181A-1	Sequence 1, Appli
43	38.4	1.9	6328	4	US-09-158-707-1	Sequence 1, Appli
44	38.4	1.9	6475	4	US-09-620-312D-325	Sequence 325, App
45	37.8	1.9	399	4	US-09-621-976-8976	Sequence 8976, Ap

ALIGNMENTS

RESULT 1  
US-08-336-345-1  
; Sequence 1, Application US/08336345  
; Patent No. 5814510  
; GENERAL INFORMATION:  
; APPLICANT: Parrieh, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,345  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30753  
; REFERENCE/DOCKET NUMBER: 7937-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Parvovirus  
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 7.9e-293;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAGGAAAAA 60

Db 273 ATGCTGGCAACCCAGTATATCTAGGAGTATATGAGGGAGTAAATGGTTAAAGAAACAT 332  
 QY 61 AGTAACAGAGAGTGTCTCATTTGTTTAAATGAAATGTTCAACTGAATGGAAGA 120  
 Db 333 GCAGAAATGAAGCATTTTCGTTGTTTAAATGTGACAACTGCTCAATGAAGAAAG 392  
 QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTGACAGGAGGAGCTGAAATCTTTACAA 180  
 Db 393 GATGTTGCTGGAAACAACTATACCAACCAATTCAAATGGAAGACTAACAATTTAAT 452  
 QY 181 CGAGGAGCGGAATACTCTTGGGACCAAGC-----GAGGACATGGAATGGGAACCA 234  
 Db 453 AGAGGAGCACAACAGCAATGATCAACCGAAGAGAAAGAAATGAGCTGGGAATCGGAA 512  
 QY 235 GTGGATGAATGACCAAAAAGCAAGTATTCATTTTGTCTTGGTTTAAATAAGTTTAA 294  
 Db 513 GTTGATAGTCTCGCCAAAAGCAAGTATCAAACTTTTGTGATTAATTAATAAATGCTTT 572  
 QY 295 TTTGGAAGTCTTAACACAAAGAAATATATTTCTGTGTGATGTTAAATGGTTTGTGCAACAT 354  
 Db 573 TTTGAAGTCTTTTCTAAATATAGAACCAATGAATGTGTTGTTTATTCACAT 632  
 QY 355 GAAATGGGAAAGACCAAGCTGGCACTGCAATGTAATAATTGGAGGAAGAGCTTTAGT 414  
 Db 633 GAATGGGAAAGATCAAGGCTGGCATTTGTCTGTTTACTTCTATAGTAAGAACTTACAA 692  
 QY 415 CAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTA 474  
 Db 693 CAAGCAACTGGTAAATGGCTACCGACACAAATGAATGATTTATGGAGTAGATGGTTGGTG 752  
 QY 475 ACAGCTGTAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATPAGCAGAA 534  
 Db 753 ACTCTTTGTTGCGTAACTTAAACCAACTGAAAGATTPAGCTCAGAGAAATPGCAGAA 812  
 QY 535 GACAATGAGTGGTACTCTTACTTATTAAGCATAGCAAAACCAAAAAGAGCTATACC 594  
 Db 813 GATAGTGAATGGTGACTATATTAACATACAGACATAAGCAAAACAAAAAGAGCTATGTT 872  
 QY 595 AAGTGTGTTCTTTTGGAAACATGATGTTCTTACTATTTTAACTAAAGAAATTAAGC 654  
 Db 873 AAAATGGTTCATTTTGGAAATATGATAGCATATTTACTTTTAAACAAAGAAATTAAGTGC 932  
 QY 655 ACTAGTCCCAAGAGACGGAGCTATTTTCTTAGCAGTGACTCTGGCTGGAATACTAAC 714  
 Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGGTGAATTTAAC 989  
 QY 715 TTTTAAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGACGCGCCA 774  
 Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGACACACCTTTTACACTGAACAAATGAACCA 1049  
 QY 775 GAAACGGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGGAGAAATTCAAACT 834  
 Db 1050 GAAACCGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGGAGAAATTCAAACT 1109  
 QY 835 AAAAAAGAGTGTCTTATTAATACTTACATTTAAAGAGCTGGTGCATTAAGAGTAACCTCA 894  
 Db 1110 AAAAGGAGTGTCAATCAATGATTTTCTGGGACTTGGTTAGTAAAGAGTAAACATCA 1169  
 QY 895 CCGAGAGCTGGATGATGATGCGAGCAGACAGTATCATTTGAATGATGGCTCAACAGGT 954  
 Db 1170 CCGAGAGCTGGATGATGATGCGAGCAGACAGTATCATTTGAATGATGGCAGCAACCGAG 1229  
 QY 955 GGAGAAACCTGCTGAAATAATAGCTAGAGATTTGTACACTAATCTAGCGCAGAACCAAA 1014  
 Db 1230 GGTGAAATCTTTTAAATAATACACTTGAATTTGTACTTTGACTTTAGCAGAGACAAA 1289  
 QY 1015 ACAGCAATTTGACTTAAATTTTAGAAAAAGCTGAAACCGAGCAAACTAACAACTTTTCACTG 1074  
 Db 1290 ACAGCAATTTGAAATTAATCTTGAAGAAAGCAGATAATACTAACTTAACTTATGATCTT 1349  
 QY 1075 CCGTACACAGAGCTGACAGATTTTGTCTTTTCACTGCGTGAACATGTTAAAGTTGCG 1134

Db 1350 GCAAAATCTAGAACCATGTCAAAATTTTAAATGACGCGATGGAATTTGATTAAGTTGT 1409  
 QY 1135 CATGCTATTTGCTGTGTTTTTAAACAGACAGAGGAGCAAAAGAAATACTGTTTTATTTTCA 1194  
 Db 1410 CACGCTATAGCATGTGTTTTTAAATAGACAAGGTGTTAAAGAAATACAGTTCTTTTTCAT 1469  
 QY 1195 GGACGAGCAGACAGGCAAAATCTATTTATTTGCAACAGCCATAGCAACAGCTTTGGCAAT 1254  
 Db 1470 GGACGAGCAAGTACAGGAAATCTATCTATGCTCAAGCCATAGCAACAGCTGTGGGTAT 1529  
 QY 1255 GTTGTGCTGCTAATAGCAGCCAAATGCTAAATTTCCATTTAATGACTGTACCAACAAAGAAC 1314  
 Db 1530 GTTGTGCTGCTAATAGCAGCCAAATGCTAAATTTCCATTTAATGACTGTACCAATTAATAAT 1589  
 QY 1315 TTGATTTGGGTAGAGAAAGCTGGTAACTTTTGACAGCAAGTAAACCAAGCTTTAAAGCCAT 1374  
 Db 1590 TTAATTTGGATTTGAAGAGCTGGTAACTTTGCTCAACAAAGTTAATCAATTTAAAGCAATC 1649  
 QY 1375 TGCTCTGCTCAAACTATTTGCAATTTGATCAAAAAGAAAGGAGCAACAGATTTGAACCA 1434  
 Db 1650 TGTTCTGGACCAACCAATTTAGAAATTTGATCAAAAAGGTAAGGAAGTAAGCAATTTGAACCA 1709  
 QY 1435 ACACGAGTCATCATCACCAACCAATCAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAAGAA 1494  
 Db 1710 ACTCAGTAATTTATGACAACTTAATGAAAATATAACAATTTGTGAGAAATTTGATGTGAGAA 1769  
 QY 1495 AGACGAGAACACATCAACCAATCAGACAGAAATGCTTTAAACATTCATCTAACACATACC 1554  
 Db 1770 AGACCTGAACATACACAACCAATTAAGAGACAGAAATGTTGAACATTTAAGTTAGTATGTAAG 1829  
 QY 1555 TTGCTCTGCTGACTTTGTTTGGTTGCAAAAATGATGGCCCATGATTTGCTTTGCTGGTTG 1614  
 Db 1830 CTTCCAGGAGACTTTGTTTGGTTGATTAAGAAAGAAATGGCCCTTTAATATGTGATGGTTA 1889  
 QY 1615 GTAAAGATGTTTCAATCTACCATGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCT 1674  
 Db 1890 GTTAAACATGTTTGAATCAACATGCTAACTATACATCATTTGGGAAAGTACCA 1949  
 QY 1675 GATGCTCAGAAAACTGGGCGGAGCCAAAGGTGCGCAACTCTTATATAATTTACTAGGTTGCG 1734  
 Db 1950 GAATGGATGAAACTGGGCGGAGCTTAAATACAAAGAGGTATAAATTCACCGGTTGC 2009  
 QY 1735 GCAGCTCACCATTACGACACCGAAAGTACGCTCTCAGCCAGAACTATATCCTACTAAT 1794  
 Db 2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAGACCAAGCTTTAAT 2066  
 QY 1795 CCACTTGCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCCTTGGAGCACACCAATACTCT 1854  
 Db 2067 CTTCTGACTCCGACGCTAGTGGACCTTGCACTGGAAACCGTGGAGTACTCCAGATAGCCT 2126  
 QY 1855 GTTGGGGCACTGCGAAGAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAGAT 1914  
 Db 2127 ATTGCAGAACTGCAAAATCAACAATCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180  
 QY 1915 GGTCACTAGCCCACTTTGTTGAGAGTTCGAGGAGGATTTGAGAGCGGTCTCGGTGGCG 1974  
 Db 2181 GTGCAAGCGAGTCCGAGCTGGTCCGAAATAGAGGAGACCTGAGAGCCATCTTACTTCT 2240  
 QY 1975 GAAACCGTTGAAGAAAGACTTTGAGCGA 2000  
 Db 2241 GAACATTTGAAGAAAGATTTTCGAGA 2266

RESULT 2

US-08-336-345-2  
 ; Sequence 2, Application US/08336345  
 ; Patent No. 5814510  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Parrish, Colin R.  
 ; APPLICANT: Gruenberg, Allen  
 ; APPLICANT: Carmichael, Ireland E.  
 ; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
 ; NUMBER OF SEQUENCES: 2



1615 GTAAAGATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTT 1674  
1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACATCATTTGGGGAAAGTACCA 1949  
1675 GATTGTGCAAAACTGGGGGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGTTGCG 1734  
1950 GAATGGGATGAAACTGGGGGAGCCTAAATACAAAGAGGTATAAATTCACCGGTTGC 2009  
1735 GCAGGCTCACCATTACGACACGAAAGTACGCTCTCAGCAGAACTATGCACCTAACT 1794  
2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCCTCAGAGTCAAGCAAGTTCTAACT 2066  
1795 CCATTTGCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCT 1854  
2067 CCTCTGACTCCGAGCTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGTAGCCT 2126  
1855 GTTGGCGGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGAT 1914  
2127 ATTGCAGAACTGCAATCAACCAATCAACCAACTTTGGGCTTACTCA-----CAAAGAC 2180  
1915 GGTCAACTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTTCGTTGCGTGG 1974  
2181 GTGCAAGCGAGTCCGACGTGGTCCGAAATAGAGGCGAGCCTGAGAGCCATCTTTACTTCT 2240  
1975 GAACCGTTGAAGAAAGACTTTCAGCGA 2000  
2241 GAACAATGGAAGAGATTTTCGAGA 2266

RESULT 3  
US-08-647-655-1  
; Sequence 1, Application US/08647655  
; Patent No. 588585  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: On Even Date Herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30,753  
; REFERENCE/DOCKET NUMBER: 7937-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Parvovirus

US-08-647-655-1

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 7.9e-293;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60  
Db 273 ATGTCTGGCAACCACTATCTGAGGAAGTTATGGAGGGAGTAAATTTGGTTAAAGAAACAT 332  
QY 61 AGTAACCGGAAGTGTCTTCATTTGTTTAAATGAAATGTTCACTGAATGGAATA 120  
Db 333 GCAGAAATGGAAGCAATTTTCGTTTGTAAATGTGCAACGTCACATAAATGGAAG 392  
QY 121 GATATCGGATGAATAGTTTACAAAAAGAGCTCGAGGAGGAGCTGGAATCTTTTACAA 180  
Db 393 GATGTTGCTGGACACACTATACCAACCAATTCAAATGAAGAGCTACATCTTTAAT 452  
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234  
Db 453 AGAGGAGCACAAACAGCAATGATCAACCCGAAAGAAAGAAATGAGCTGCGGAATCGGA 512  
QY 235 GTGGATGAATGACCAAAAAAGCAAGTATTCATTTTGTATCTTTGGTTAAANAATGTTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAAACTTTTGATGCAATTAATTAANAATGCTCT 572  
QY 295 TTTGAAGTGTCTTAACACAAAGAAATATATTTCTGTGTGATGTTAAATTTGGTTTGTGCAACAT 354  
Db 573 TTTGAAGTGTCTTTTCTTAAANAATAGAACCAATGATGTGTTTGTGTTTATTCACAT 632  
QY 355 GAATGGGAAAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGT 414  
Db 633 GAATGGGAAAAAGATCAAGCTGGCATTTGTCTGTTTACTTTTATAGTAAGAACTTACAA 692  
QY 415 CAACTCAAGGAAATGTGGAGAGAGCACTAAATGTTTACTGGAGCAGATGTTGGTA 474  
Db 693 CAAGCAACTGTGTAATGGCTACGACAGACAAATGAAATATGTTATGGAGTAGATGTTGGTG 752  
QY 475 ACAGCTGTAATGTGCAACTTAACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAA 534  
Db 753 ACTTTTGTTCGTAAACCTTAACCACTGAAAGATTTAGCTCAGAGAAATTCAGAA 812  
QY 535 GACAAATGAGTGGTTTACTCTACTTATATAGCATTAAGCAACCAACCAAGCACTATACC 594  
Db 813 GATAGTGAATGGTGCATATATTAACATACAGACATAGCAACCAACCAAGCACTATGTT 872  
QY 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGC 654  
Db 873 AAAATGGTTTCATTTTGGAAATATGATAGCATATTTACTTTTAAACAAAGAAAAATTTGC 932  
QY 655 ACTAGTCCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAAC 714  
Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTCGTTGGGAAATTTAAC 989  
QY 715 TTTTAAAAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAGCACACTTTTACACTGAACAAATGAAACCA 1049  
QY 775 GAAACGGTTGAAACCAAGTAAACCACTGCGGAGGAAACTAAGCGGCGGAGAAATTCAACT 834  
Db 1050 GAAACGGTTGAAACCAAGTAAACCACTGCGGAGGAAACCAAGCGGCGGAGAAATTCAACT 1109  
QY 835 AAAAAGAGAGTCTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCA 894  
Db 1110 AAAAGAGAGTGTCAATCAATGTTACTTTGCGGAGCTTGTGTTAGTAAAGAGTAACTCA 1169  
QY 895 CCAGAGGACTGGATGATGATGCGAGCAGACAGTTACATTTGAAATGATGCTCAACAGGT 954  
Db 1170 CTTGAAGCTGGATGATGTTTACAAACAGATAGTTATATTGAAATGATGACCAACCCAGGA 1229  
QY 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCGAGCAACCA 1014  
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTTGTACTTTGACTTTTGAAGAGCAAAA 1289



Qy	415	CAAGCTCAAGGGAATGGTGGAGAAGGCACATAAATGCTTTACTGGACGAGATGGTTGGTA	474
Db	693	CAAGCAACTGGTAAATGGCTACGCAGACAAATGAATATGTATTGGAGTAGATGGTTGGTG	752
Qy	475	ACAGCCTGTAATGTGGCAACTTAACACACGAGTGAAGAATTTAACTAAGAGAAATAGCAGAA	534
Db	753	ACTCTTTGTTTCGGTAAACTTAAACCAACCTGAAAAGATTAAGCTCAGAGAAATTGCAGAA	812
Qy	535	GACAAATGAGTGGGTACTCTACTTTACTTATAAGCATAAAGCAAAACCAAAAAGAGACTATACC	594
Db	813	GATAGTGAATGGGTGACTATATTAAACAATACGACATTAAGCAAAACAAAAGAGACTATGTT	872
Qy	595	AAGTGTGTTCTTTTGGAAACATGATTTGCTTACTATTTTTTTAACTAAAAAGAAAAATAAGC	654
Db	873	AAAAATGGTTCATTTTGGAAATATGATAGCATATTACTTTTTTAAACAAAGAAAAAAATGTC	932
Qy	655	ACTAGTCCACCAAGAGACGGAGGCTATTTCTTTAGCAGTGACTCTGCTGGAAAACTTAAC	714
Db	933	CACAT---GACAAAAGAAAGTGGCTATTTTTTAAAGTACTGATTTCTGGTTGGAAATTTTAA	989
Qy	715	TTTTTTAAAGAAAGCGAGCGGCATCTAGTGAGCAAACTATACACTGATGATCATGCGGGCCA	774
Db	990	TTTTATGAAGTATCAAGACAGACAAATTTGTTCAGCACACTTTACACTGACAAATGAACCA	1049
Qy	775	GAACCGGTTGAAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACT	834
Db	1050	GAACCGGTTGAAACCAACAGTACGACAGCACAGGAAACAAAGCGCGGAGAAATTCAAACT	1109
Qy	835	AAAAAGAGTTCTATTATTAACACTACTTTAAAGAGCTGGTGCATTAAGAGTAACCTCA	894
Db	1110	AAAAAGGAAGTGTCAATCAAAATGTACTTTGCGGACTTTGGTTAGTAAAGAGTAACATCA	1169
Qy	895	CCAGAGACTGGATGATGATGCAGCCACAGATTTACATTGAAATGATGGCTCAACACAGT	954
Db	1170	CCTGAAGACTGGATGATGTTACACACAGATGTTATATTGAATGATGGCACACACAGGA	1229
Qy	955	GGAGAAAACTGTGAAAAATACCGCTAGAGATTTGTTACATTAECTCTAGCCAGAACCAAA	1014
Db	1230	GGTGAAAAATCTTTTAAAAAATACACTTTGAAATTTGTACTTTTGACTTTAGCAAGAACAAA	1289
Qy	1015	ACAGCATTGCACTTAATTTTAGAAAAAGCTGAACACGACGAACTAACCAACTTTTCACTG	1074
Db	1290	ACAGCATTGGAATTAATACTTTGAAAAAGCAGATAATCTAAACTTAATCTTTTGATCTTT	1349
Qy	1075	CCTGACACAAGAGCCTGCAGAAATTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTTC	1134
Db	1350	GCAAAATCTAGAACATGTCAAATTTTTAGAATGCAACGGATGGAATTCGAATTAAGTTTGT	1409
Qy	1135	CATGCTATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCAT	1194
Db	1410	CACGCTATAGCATGTGTTTTTAAATAGACAAGGTGTTAAAAAGAAATACAGTTCTTTTTCAT	1469
Qy	1195	GGACACGACGACACAGGCAAAATCTATTATTGSCAACAGCCATAGCACAAAGCAGTTGGCAAT	1254
Db	1470	GGACACGACAGTACAGGAAATCTATCTATGCTCAAGCCATAGCACAAAGCTGTGGGTAAAT	1529
Qy	1255	GTTGGTTGCTATAATGCAGCCAAATGTTAAACTTTCCATTTAATGACTGTACCAACAAGAAC	1314
Db	1530	GTTGGTTGTTATAATGCAGCAAAATGTTAAATTTTCCATTTAATGACTGTACCAATAAAAAAT	1589
Qy	1315	TTGATTTTGGGTAGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACACGAGTTTAAAGCAATT	1374
Db	1590	TTAAATTTGGATTTGAAGAAGCTGGTAACTTTTGGTCAACAGTTAATCAATTTTAAAGCAATC	1649
Qy	1375	TGCTCTGGTCAAACTATTTCGCATTTTGATCAAAAAGGAAAAAGCGACCAACAGATTTGAACCA	1434
Db	1650	TGTTCTGGACAAAACAAATTAGAAATTTGATCAAAAAGGTAAAGGAGTAGCAAAATTTGAACCA	1709
Qy	1435	ACACAGTCTATGACACAAAATGAGAAACATTAACGTGGTTCAGAAATAGCGTCGCAAGAA	1494
Db	1710	ACTCCAGTAAATTATGACAACTTAATCAAAATATAACAATTTGTGAGAAATGGAATGTGAAGAA	1769

Qy	1495	AGACCAGAACACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTCACTTAAACACATACC	1554
Db	1770	AGACCTGAAACATACACAAACCAATTAAGAGACAGAAATGTTGAACATTAAGTTAGTATGTAAG	1829
Qy	1555	TTGCGCTGGTCACATTTGGTTGGTTGTGACAAAATGAATGGCCCATGAATTTGTGCTTGCTTG	1614
Db	1930	CTTCCAGGAGACATTTGGTTGGTTGTATTAAGAGAAATGGCCTTTAATAATGTGCATGGTTA	1989
Qy	1615	GTTAAGAATGGTTACCAATCTACATGGCCAAAGCTACTGTGCTCTAAATGTGGGCAAAAGTTCCT	1674
Db	1890	GTTAAAACATGGTTTGAATCAACCATGGCTTAACATATACACATCATTTTGGGGAAAAGTACCA	1949
Qy	1675	GATTTGGTCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTTAGTGTTCG	1734
Db	1950	GAATGGGATGAAAATCTGGGCGGAGCCTTAAATAACAGAGGTAATAATTCACCAAGTTGC	2009
Qy	1735	GCACGCTCACCATTTACGACACCCGAAAGTAGTACGCCCTCTCAGCCAGAACTATGACACTAACT	1794
Db	2010	AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCTCAGAGTCAGAGCCAAAGTTCTAACT	2066
Qy	1795	CCACTTGCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATACTCCT	1854
Db	2067	CCTCTGACTCGGACGTAGTGGACCTTTGCACTGGAAACCGTGGAGTACTCCAGATACGCCT	2126
Qy	1855	GTTTCGGGCACTGCAGAAACCCAGAAACACTGGGGAAAGCTGGTTCCAAAGAGCTGCCAAGAT	1914
Db	2127	ATTGCAAAAACCTGCAAAATCAACAATCAAAACCAACTTGGCGCTTACTCA-----CAAAGAC	2180
Qy	1915	GGTCACTAGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTTCCGGTGGC	1974
Db	2181	GTGCACGCGAGTCCGACGTGGTCCGAATATAGAGGCGAGCCTGAGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTTGAAGAAAAGACTTCAGCGA	2000
Db	2241	GAACAAATTGGAAAGAAATTTTCAGA	2266

RESULT 5  
US-08-254-358-1  
; Sequence 1, Application US/08254358  
; Patent No. 5658785  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; METHOD OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/254,358  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5658785and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs







CURRENT APPLICATION NUMBER: US/09/807,802A  
CURRENT FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: US 60/107,114  
PRIOR FILING DATE: 1998-11-05  
PRIOR APPLICATION NUMBER: PCT/US99/25694  
PRIOR FILING DATE: 1999-11-02  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18  
LENGTH: 4681  
TYPE: DNA  
ORGANISM: AAV-2  
US-09-807-802A-18

Query Match 5.4%; Score 109.4; DB 4; Length 4681;  
Best Local Similarity 53.7%; Pred. No. 2.9e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCNAAAGAACTACTGTTTATTTTCATGACCCAGCAGCAGGCAAAATCTATTATTGCA 1227  
DB 1293 GGCNAAAGAACTACTGTTTATTTTCATGACCCAGCAGCAGGCAAAATCTATTATTGCA 1352  
QY 1228 CAAGCCATAGCAACAGCAGTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTTT 1287  
DB 1353 GAGGCATAGCAACAGCAGTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTTT 1412  
QY 1288 CCATTTAATGACTGTATACCAAGAACTTTGATTTGGGTAGAAAGCTCGTAACTTTGGA 1347  
DB 1413 CCCTTCAACGACTGTGTCACAAAGATGTTGTTGGGTAGAAAGCTCGTAACTTTGGA 1472  
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGATTGATCAAAAA 1407  
DB 1473 AAGGTCTGGAGTCGCGCAAAAGCCATTTCTCGGAGGAAGCAAGTGGCGGTGACCAAGAA 1532  
QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACAGCTATCATGACCAACAAATGAGAACATT 1467  
DB 1533 TCGAAGTCTCGGCCCAAGATAGACCCGATCTCCGTGATGTCACCTCCAAACCAACATG 1592  
QY 1468 ACAGTGGTCAGAAATAGGCTGCGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527  
DB 1593 TCGCGCGTATTGACGGGAACCTCAACGACCTTCGAAACACAGCAGCGTTGCAAGACCGG 1652  
QY 1528 ATGCTTAACTTCACTTACACATACCTTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1587  
DB 1653 ATGTTCAAAATTTGAATCTACCCCGCTCTGGATCATGACTTTTGGGAAGGTCCCAAGCAG 1712  
QY 1588 GAA 1590  
DB 1713 GAA 1715

## RESULT 10

US-08-331-384-2/c  
Sequence 2, Application US/08331384  
Patent No. 5856152  
GENERAL INFORMATION:  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331,384  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: UPNG1149USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLSCULE TYPE: cDNA  
US-08-331-384-2

Query Match 5.4%; Score 109.4; DB 2; Length 4910;  
Best Local Similarity 53.7%; Pred. No. 2.9e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCNAAAGAACTACTGTTTATTTTCATGACCCAGCAGCAGGCAAAATCTATTATTGCA 1227  
DB 1544 GGCNAAAGAACTACTGTTTATTTTCATGACCCAGCAGCAGGCAAAATCTATTATTGCA 1485  
QY 1228 CAAGCCATAGCAACAGCAGTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTTT 1287  
DB 1484 GAGGCCATAGCAACAGCAGTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTTT 1425  
QY 1288 CCATTTAATGACTGTATACCAAGAACTTTGATTTGGGTAGAAAGCTCGTAACTTTGGA 1347  
DB 1424 CCCTTCAACGACTGTGTCACAAAGATGTTGTTGGGTAGAAAGCTCGTAACTTTGGA 1365  
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGATTGATCAAAAA 1407  
DB 1364 AAGGTCTGGAGTCGCGCAAAAGCCATTTCTCGGAGGAAGCAAGTGGCGGTGACCAAGAA 1305  
QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACAGCTATCATGACCAACAAATGAGAACATT 1467  
DB 1304 TCGAAGTCTCGGCCCAAGATAGACCCGATCTCCGTGATGTCACCTCCAAACCAACATG 1245  
QY 1468 ACAGTGGTCAGAAATAGGCTGCGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527  
DB 1244 TCGCGCGTATTGACGGGAACCTCAACGACCTTCGAAACACAGCAGCGTTGCAAGACCGG 1185  
QY 1528 ATGCTTAACTTCACTTACACATACCTTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1587  
DB 1184 ATGTTCAAAATTTGAATCTACCCCGCTCTGGATCATGACTTTTGGGAAGGTCCCAAGCAG 1125  
QY 1588 GAA 1590  
DB 1124 GAA 1122

## RESULT 11

US-08-836-087-2/c  
Sequence 2, Application US/08836087  
Patent No. 5871982  
GENERAL INFORMATION:  
APPLICANT: Trustees of University of Pennsylvania  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457



Db 1184 ATGTTCAAAATTTGAACCTACCCCGCTGCTGGATCATGACTTTTGGGAAGGTCACCAAGCAG 1125

QY 1588 GAA 1590  
|||

Db 1124 GAA 1122

RESULT 13  
US-09-546-738-2/c  
; Sequence 2, Application US/09546738  
; Patent No. 6387368  
; GENERAL INFORMATION:  
; APPLICANT: Trustees of University of Pennsylvania  
; Wilson, James M.  
; Kelley, William M.  
; Fisher, Krishna J.  
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
; Methods of Use Thereof  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, PO Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/546,738  
; FILING DATE: 11-Apr-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/246,320  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: GNPVN.007PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4910 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-546-738-2

Query Match 5.4%; Score 109.4; DB 3; Length 4910;  
Best Local Similarity 53.7%; Pred. No. 2.9e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GCGAAAGAAATCTGTTTATTTATGACAGCCAGCAGGCAAAATCTATTATGCA 1227

Db 1544 GCGAAGAGAAACACCATCTGGCTGTTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 1485

QY 1228 CAAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTTT 1287

Db 1484 GAGGCCATAGCCACACTGTGCCCTTCTAGGGTGGTAACTGGACCAATGAGAATTT 1425

QY 1288 CCATTTAATGACTGTACCAACAGAACTTTGATTTGGGTAGAGAAGCTGGTAACTTTGA 1347

Db 1424 CCCTTCAACGACTGTGTGCAAGAAGTGGTGATCTGGTGGGAGGAGGAGATGACCGCC 1365

QY 1348 GAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGCAATTGATCAAAAA 1407

Db 1364 AAGTCTGTGAGTCGGCCAAAGCCATTTCTCGGAGGAAGAGGTCGCCGTGGACAGAAA 1305

QY 1408 GGAAGAGGAGCAAAACAGATTGTAACCAACACCAAGTCTCATCATGACCAAAATGAGAACATT 1467

Db 1304 TGCAAGTCTCTCGGCCCCAGATAGACCCGACTCCCGTGTATCGTCACTCCAACACCAATG 1245

QY 1468 ACAGTGTGTCAGAAATAGGCTGCGAAGAAAGACAGACACACTCAACCAATCAGAGACAGA 1527

Db 1244 TGGCCGCTGATTGACGGGAACCTCAACGACCTTCGGAACACACGAGCGGTTTGAAGACCGG 1185

QY 1528 ATGCTTAAACATTTCACTTAACACATACCTTGCCTGGTGACTTTTGGTTTGGTTCACAAAAT 1587

Db 1184 ATGTTCAAAATTTGAACCTACCCCGCTGCTGGATCATGACTTTTGGGAAGGTCACCAAGCAG 1125

QY 1588 GAA 1590  
|||

Db 1124 GAA 1122

RESULT 14  
US-09-438-268-1/c  
; Sequence 1, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Radinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 7214  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-1

Query Match 5.4%; Score 109.4; DB 4; Length 7214;  
Best Local Similarity 53.7%; Pred. No. 3.6e-21;  
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GCGAAAGAAATCTGTTTATTTATGACAGCCAGCAGGCAAAATCTATTATGCA 1227

Db 3899 GCGAAGAGAAACACCATCTGGCTGTTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 3840

QY 1228 CAAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTTT 1287

Db 3839 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGGCTAACTGGACCAATGAGAATTTT 3780

QY 1288 CCATTTAATGACTGTACCAACAGAACTTTGATTTGGGTAGAGAAGCTGGTAACTTTGA 1347

Db 3779 CCCTTCAACGACTGTGTGCAAGAATGGTGATCTGGTGGAGGAGGAGATGACCGCC 3720

QY 1348 GAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGCAATTGATCAAAAA 1407

Db 3719 AAGGTCTGGAGTCGGCCAAAGCCATTTCTCGGAGGAAGCAAGGTGCGCGTGGACAGAAA 3660

QY 1408 GGAAGAGGAGCAAAACAGATTGAAACCAACACCAAGTCACTATGACCAACATGAGACATT 1467

Db 3659 TCAAGTCTCTCGGCCCCAGATAGACCCGACTCCCGTGTATCGTCACTCCAACACCAATG 3600

QY 1468 ACAGTGTGTCAGAAATAGGCTGCGAAGAAAGACAGACACACTCAACCAATCAGAGACAGA 1527

Db 3599 TCGCCGCTGATTGACGGGAACCTCAACGACCTTCGAACACACGAGCGGTTTGAAGACCGG 3540

QY 1528 ATGCTTAAACATTTCACTTAACACATACCTTGCCTGGTGAATTTGGTTTGGTTCACAAAAT 1587



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds  
(without alignments)  
10821.780 Million cell updates/sec

Title: US-10-069-056-8  
Perfect score: 2019  
Sequence: 1 atggctggaatgctactc.....agccgctgaacttgactaa 2019

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications\_NA.\*
- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
  - 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
  - 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
  - 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
  - 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
  - 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
  - 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
  - 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
  - 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
  - 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
  - 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
  - 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
  - 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
  - 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
  - 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
  - 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
  - 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
  - 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	US-10-647-111A-1	Sequence 1, Appli
2	941.4	46.6	5075	US-10-361-002-3	Sequence 3, Appli
3	941.4	46.6	5075	US-10-361-004-3	Sequence 3, Appli
4	307.4	15.2	374	US-10-466-894-1276	Sequence 1276, Ap
5	285.8	14.2	421	US-10-466-894-430	Sequence 430, App
6	284.8	14.1	423	US-10-466-894-375	Sequence 375, App
7	284.8	14.1	464	US-10-466-894-370	Sequence 370, App
8	284.8	14.1	473	US-10-466-894-367	Sequence 367, App
9	284.8	14.1	486	US-10-466-894-358	Sequence 358, App
10	284.8	14.1	491	US-10-466-894-428	Sequence 428, App
11	284.2	14.1	420	US-10-466-894-438	Sequence 438, App
12	283.8	14.1	343	US-10-466-894-388	Sequence 388, App

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, Ap
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	117	5.8	1386	9	US-09-792-630-32	Sequence 32, Appl
40	117	5.8	1386	10	US-09-953-351-32	Sequence 32, Appl
41	117	5.8	1386	13	US-10-080-376-32	Sequence 38, Appl
42	117	5.8	1386	14	US-10-082-671-38	Sequence 38, Appl
43	117	5.8	1386	14	US-10-097-100-32	Sequence 32, Appl
44	117	5.8	1386	15	US-10-023-208-32	Sequence 32, Appl
45	115.4	5.7	1884	9	US-09-792-630-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-647-111A-1  
; Sequence 1, Application US/10647111A  
; Publication No. US20040209240A1  
; GENERAL INFORMATION:  
; APPLICANT: ICGO, RICHARD  
; APPLICANT: MALERBA, MADDALENA  
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES  
; FILE REFERENCE: 604-691  
; CURRENT APPLICATION NUMBER: US/10/647,111A  
; CURRENT FILING DATE: 2003-08-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 5121  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1  
; OTHER INFORMATION: with promoter P4 and left hairpin from WVM (phH1)  
US-10-647-111A-1

Query Match	87.7%	Score	1771;	DB	18;	Length	5121;
Best Local Similarity	92.3%	Pred. No.	0;				
Matches	1864;	Conservative	0;	Mismatches	155;	Indels	0;
Gaps	0;						
QY	1	ATGGCTGGAATGCTTACTCTGATGAACCTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
DB	282	ATGGCTGGAATGCTTACTCTGATGAACCTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
QY	61	AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120				
DB	342	AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	401				
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTTACAA	180				

402	Db		GATATCGGATGGAATAGTTTACAAAAAAGAGCTGCGAGGAGCGAGCTGAAATCTTTACAA	461
181	Qy		CGAGGAGCGAAACTACTTTGGGACAAAGCGAGGACATGGAATGGAAACACACAGTGGAT	240
462	Db		CGAGGAGCGGAAACTACTTTGGGACCAAAGCGAGGACATGGAATGGAAACACACAGTGGAT	521
241	Qy		GAAATGACCAAAAGCAAGTATTTCATTTTGTATCTTTTGGTTTAAAAATGTTTATTGAA	300
522	Db		GAAATGACCAAAAGCAAGTATTTCATTTTGTATCTTTTGGTTTAAAAATGTTTATTGAA	581
301	Qy		GTGCTTAAACAAGAAATATATTTTCTGTGTGATGTTTAAATTTGGTTTGTGCAACATGAATGG	360
582	Db		GTGCTTAAACAAGAAATATATTTTCTGTGTGATGTTTAAATTTGGTTTGTGCAACATGAATGG	641
361	Qy		GGAAAAACAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTGTAGTCAAGCT	420
642	Db		GGAAAAACAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTGTAGTCAAGCT	701
421	Qy		CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGTGTAAACAGCC	480
702	Db		CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGTGTGTAAACAGCC	761
481	Qy		TGTAATGTGCAACTAAACACCAAGCTGGAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT	540
762	Db		TGTAATGTGCAACTAAACACCAAGCTGGAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT	821
541	Qy		GAGTGGGTACTCTACTTACTTATAACATATAGCAATACCAACCAAAAAGACTATACCAAGTCT	600
822	Db		GAGTGGGTACTCTACTTACTTATAACATATAGCAATACCAACCAAAAAGACTATACCAAGTCT	881
601	Qy		GTTCCTTTTGGAAACATGATTGCTTACTATTTTAAAACTAAAAAGAAAAATAGACACTAGT	660
882	Db		GTTCCTTTTGGAAACATGATTGCTTACTATTTTAAAACTAAAAAGAAAAATAGACACTAGT	941
661	Qy		CCACAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTTTA	720
942	Db		CCACAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTTTA	1001
721	Qy		AAAGAGGCGAGCGGCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG	780
1002	Db		AAAGAGGCGAGCGGCATCTAGTGAGCAAACTATAGTGAGCAAACTGTTACTGATGAGATGAAAC	1061
781	Qy		GTGAAAAACACAGTAAACCACTGCGGACGAAACTAAGCGCGCAGAAATTCAACTTAAAAA	840
1062	Db		GTGAGAACACAGTGAACCACTGCAAGAGCTAAGCGCGCAGAAATTCAACTTAGAGAG	1121
841	Qy		GAAGTTCCTATTAAAACTACACTTAAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	900
1122	Db		GAGGTCTCGATTAAAAACCACTCAAGAGTTGGTACATAAAAGAGTAACCTCACAGAA	1181
901	Qy		GACTGGATGATGATGACGACGAGTTCATTGAAATGATGGCTCAACAGGTGGAGAA	960
1182	Db		GACTGGATGATGATGACGACGAGTTCATTGAAATGATGGCTCAACAGGTGGAGAA	1241
961	Qy		AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTAGCCAGAACCAAAACAGCA	1020
1242	Db		AACCTGCTGAAAAATACGCTAGAGATCTGTACACTGACTCTAGCAAGAACCAAAACAGCC	1301
1021	Qy		TTTGACTTAAATTTTAGAAAAAGCTGAAACACAGCAAACTAAACCACTTTTCACTGCTGAC	1080
1302	Db		TTTGACTTGAATTTCTGAAAAAGCTGAAACACAGCAAACTAGCCAACTTTTCCATGGCTAGC	1361
1081	Qy		ACAAGAGCCTGAGAAATTTTCTGTTTCAATGCTGGAATATGTTTAAAAAGTTTGCATGCT	1140
1362	Db		ACCAGAACCTGTAGAATCTTTGCTGAGCAAGCTGGAACTATATTAAAGTCTGCAATGCC	1421
1141	Qy		ATTGTGCTGTTTAAACACAGACAGGAGCAAAAGAAATACGTGTTTATTATTCATGAGACCA	1200
1422	Db		ATCTGTGTGCTGTAATAGACAAGGAGCAAAAGGAAACACTGTGCTCTTTTACGAGACCA	1481
1201	Qy		GCAGACACAGGCAATCTATTATTGACAAAGCCATAGACAAGAGTGTGGCAATGTTGGT	1260

Db	1482	GCCAGCACAGGCAAACTCTATTATTCGCAAGCCATAGCAACAAGCAGTGGTAAATGTTGGT	1541
Qy	1261	TGCTAATAATCGACGCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAAGAAGAACTTGATT	1320
Db	1542	TGTTTACAAATGCTGCCAATGTGAACATTTTCATTTAATGACTGTACCAACAAGAAGAACTTGATT	1601
Qy	1321	TGGGTAGAGAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTTGCTCT	1380
Db	1602	TGGGTGGAAGAGCTGGTAACTTTTGGCCAGCAAGTAAACCAATTTCAAAGCTATTTGTGTTCT	1661
Qy	1381	GGTCAAACTATTGGCATTTGATCAAAAGGAAGGAGCAAAACAGATTGAACCAACACCA	1440
Db	1662	GGCCAAACCAATGACATTTGATCAAAAGGAAGGAGCAAAACAGATTGAACCAACACCA	1721
Qy	1441	GTCAATCATGACCAAAATGAGAACTTATACATGTGGTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Db	1722	GTTATATTAGCCACCAACGAGAACATTTACCGTTGGTTAGAAATAGGCTGTGAGGAAGACCA	1781
Qy	1501	GAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCATCTACACATACCTTGGCCT	1560
Db	1782	GAAACACACTCAACCAATCAGAGACAGAAATGCTCAACATTCACCTGACACAGTACACCT	1841
Qy	1561	GGTGACTTTGGTTTGGTTGTGCAAAATATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
Db	1842	GGTGACTTTGGTTTGGTTGGATTAACACAGATAGGCTCTGATCTGTGCTTGGTTGGTGAAG	1901
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTTCCTGATTGG	1680
Db	1902	AATGGTTACCAATCTACCATGGCTTGTATTCTGTGCTAAATGGGGCAAAAGTTTCCTGATTGG	1961
Qy	1681	TCAGAAATCTGGCGGAGCCAAAGGTGCCAACTCCCTATAAATTTACTPAGGTTTCGGCACGC	1740
Db	1962	TCAGAGAGCTGGCGGAGCCGAAGCTAGACACTCCTATAAATTCGCTPAGGTTCAATGGCG	2021
Qy	1741	TCACCATTTACGACACACCGAAAGTAGCGCTCTACGCCAGAACTATGCACTPAATCTCACTT	1800
Db	2022	TCACCATCTCTGACTTCGAGAGTAGCGCTCTACGCCAAAACTACGCTCTTACTCCACTT	2081
Qy	1801	GCATCCGATCTCGAGACCTCGGCTTTAGAGCCTTGGAGCACCAAAATACTCCTGTTGGC	1860
Db	2082	GCATCCGACCTTTCGCGACTTAGCTCTAGAGCCTTGGAGCACCAAAATACTCCTGTTGGC	2141
Qy	1861	GGCATGCGAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA	1920
Db	2142	GGCATCTGACGAGAGCCAAACACTGGGAGGCTGGTTCCACAGCCTGCCAAGGTGCTCAA	2201
Qy	1921	CTGAGCCCAACTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTCTCGGTCCGGAACCG	1980
Db	2202	CGGAGCCCAACTGGTCCGAGATCGAGGCGGATTTGAGAGCTTGCTTCAGTCAAGAACAG	2261
Qy	1981	TTGAGAAAGACTTCAGCGAGCCGCTGAACTTGGACTAA	2019
Db	2262	TTTGAGAGCGACTTCACGAGAGCTGACCTTGGACTAA	2300

```

RESULT 2
US-10-361-002-3
; Sequence 3, Application US/10361002
; Publication No. US20040170954A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gilmeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Armistead, David
; TITLE OF INVENTION: Pathogen Inactivation Assay
; FILE REFERENCE: CI-0043
; CURRENT APPLICATION NUMBER: US/10/361,002
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 5075

```



TYPE: DNA  
ORGANISM: Porcine parvovirus  
US-10-361-002-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;  
Best Local Similarity 70.4%; Pred. No. 2.1e-237;  
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

QY	4	GCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAGT	63
Db	298	GGGGGAAACACTTACTCGGAAGAGGTACTAAAGCTACCAACTGGCTTCAAGATATGCT	357
QY	64	ACCAGGAGTGTCTCAATTTTAAATGAAATGTTCAACTGAATGGAAGAT	123
Db	358	CAAAAGAGACATCTCTTATGATTTAAACACAAAGAGTCAATCTAAATGGAAAGAA	417
QY	124	ATCGGATGGAATAGTTTACAAAAGAGCTGCAGGAGGAGCTGAAATCTTTTACAACGA	183
Db	418	ATTGCTTGGATTAACCAACAAAGATACACAGATGCGGAATGATTAACCTACAAAGA	477
QY	184	GGAGCGGAAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAAACCAAGTGGATGA	243
Db	478	GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCGAAGCAGC	537
QY	244	ATGACCAAAAGCAAGTATTCATTTTGAATCTTGGTTAAAGATGTTTATTTGAAGTG	303
Db	538	CTCAGAAAACGGCAAGTACTGATTTTGAATCTCTTGTAAAGATGCTCTTTGAAGGT	597
QY	304	CTTAAACAAAAGATATATTTCTGCTGATGTTAATTTGGTTTGTGCAACATGAATGGGA	363
Db	598	ATATTGCAAAAGAACCTAAGTCAAGTGAATGCTGCTTCTTCTTAAAGATGCTCTTTGAAGGT	657
QY	364	AAAGACCAAGGCTGGCACTGCCATGTAATTTGGAGGAAAGGACTTTTGTCAAGCTCAA	423
Db	658	CAAGTACTGGCTTACTCACTGCCATGTAATTTGGAGGAAAGGCTTTTCAACCAAGCAATG	717
QY	424	GGAAATGTTGGAAGGCACTAATTTTACTGGAGCAGATGGTTGTTAAACAGCCTGT	483
Db	718	GGAAATGTTGGAAGGCACTAATTTTACTGGAGTGAATTTTAAATGGAATGGAATGGAATG	777
QY	484	ATGTGCACTAACACGCTGAAGAAATTAAGTAAAGAAATAGCAGAAAGCAATGAG	543
Db	778	AAAGTACCTTAACACGCTGAAGAAATTAAGTAAAGAAATAGCAGAAAGCAATGAG	837
QY	544	TGGGTACTCTACTTATTAAGCAATAAGCAACCAAAAGGACTATATACCAAGTGTGT	603
Db	838	TGGGTATCGCTACTTACTTACCTACCTACCAACCAAACTTAAAGCAATATACAAATGACT	897
QY	604	CTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAATAGCAGTACTAGTCCA	663
Db	898	CAATTTGGAAATATGATTTGCTTACTTCTTAAATTAAGAAAGAAAGCAACT	951
QY	664	CCAAGAGCGGAGGCTATTTCTTACAGTACTCTGGCTGGGAAAGTAACTTTTAAAG	723
Db	952	GAAAGAGCATGATATTTATCTCAGCTCAGATTTCTGGCTTCTAGCAAAATTTCTTAAAG	1011
QY	724	GAAAGGAGCGGCTTACTGATGAGCAAACTATACATGATGATGCGGCGGAGAAACGGTT	783
Db	1012	GAAAGGAGGAGACACTTAGTCACTACCTATTTACTGAGCAATTAACCTGAACTGTG	1071
QY	784	GAAACACAGTAACTCTGGGAGGAACTAAGCGGCGGAGATTCAAATCTTAAAGAA	843
Db	1072	GAAACAAACGGTTACTACAGCTCAGGAAGCAAAAGAGGAGGAGATACAAACAAAGAA	1131
QY	844	GTTTCTTAAAGTACATTAAGAGCTGGTGCATTAAGAGTAACTCACCAGAGAC	903
Db	1132	GTAAGCATAAATGCAATAGAGACTTGGTTTAAATAAGATGTTACTAGCATAGAAGAC	1191
QY	904	TGATGATGATGAGCAGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATG	963
Db	1192	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1251
QY	964	CTGCTGAAAGATACGCTAGAGATTTGTACACTTCTAGCCAGAACCAAAACAGCATTT	1023

RESULT 3  
US-10-361-004-3  
; Sequence 3, Application US/10361004  
; Publication No. US20040170981A1  
; GENERAL INFORMATION:  
; APPLICANT: Clearant, Inc.  
; APPLICANT: McKenney, Keith  
; APPLICANT: Gillmeister, Lidja  
; APPLICANT: Marlowe, Kristina  
; APPLICANT: Armistead, David  
; TITLE OF INVENTION: Real-time Polymerase Chain Reaction Using Large Target Amplicons  
; FILE REFERENCE: CI-0042  
; CURRENT APPLICATION NUMBER: US/10/361,004

Db	1252	TTAATCAAAATATACACTAGAAATTAACAATCTTCTTCTAGCAAGAAACAAACAGCATAT	1311
QY	1024	GACTTAAATTTTAAAGAAAGCTGAAACCAAGCAACCAACCAACCAACCAACCAACCAAC	1083
Db	1312	GACTTAAATTTTAAAGAAAGCTGAAACCAAGCAACCAACCAACCAACCAACCAACCAAC	1371
QY	1084	AGAGCCTCAGAAATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1143
Db	1372	AGAACATGTAATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1431
QY	1144	TGCTGTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1203
Db	1432	ACTTGTGTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1491
QY	1204	AGCAGGCAAAATCTAATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1263
Db	1492	TCAACAGGAAAGATATATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1551
QY	1264	TATATGAGCAATGTAATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1323
Db	1552	TACAAATGCAAGCAATGTAATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1611
QY	1324	GTAGAGAGCTGTTAACTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1383
Db	1612	ATTGAAGAGCAAGCAATTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTG	1671
QY	1384	CAAACTATTTCGCAATGATCAAAAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1443
Db	1672	CAACCAATTTAGATTTGACCAAAAGGTAAGGAAAGGTAAGGAAAGGTAAGGAAAGGTAAGG	1731
QY	1444	ATCATGACCAAAATGAGAACTTTACAGTGGTCAAGATAGGCTGCGAAGAAAGCAAGAA	1503
Db	1732	ATAATGACTTACAAATGAGACATACTAAAGTTAGAAATAGGATGCGGAGGAAAGCAAGAA	1791
QY	1504	CACACTCAACCAATCAGAGCAGAGATGCTTAACTTCACTTCACTTCACTTCACTTCACTTCACT	1563
Db	1792	CATACACCAATTAAGAGCAGAAATGTTTAAACATAAATTAACCAACCAACCAACCAACCAAC	1851
QY	1564	GACTTGTGTTTGGTTGACAAATGAATGGCCCATGATTTGCTTTGCTTTGCTTTGCTTTGCTTTG	1623
Db	1852	GATTTTGGATTTTAGAGAACTGAAATGCGCACTAATATGCTTTGCTTTGCTTTGCTTTGCTTTG	1911
QY	1624	GTTTACCAATTTACATGCGCAAGCTACTGTCTTAAATGCGGCAAGTCTCTGATGCTCA	1683
Db	1912	GGTTACCAAGCAAAATGCTAGCTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG	1971
QY	1684	GAAACTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1743
Db	1972	GAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2028
QY	1744	CCATTCACGACACCGAAAGTACGCTCTCTGAGCGAGAACTATGCACTTAACTTCACTTTGCA	1803
Db	2028	ATTTCACATCAGTGAAGAACTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2088
QY	1804	TCGGATCTCGA---GGACTGGCTTTAGAGCTTTAGAGCTTTGGAGCAGCAGCAAAATCTCC	1853
Db	2089	GAGGACCTGGAATTTAGCTTTAGCTTTGGAGCTTTGGAGCTTTGGAGCTTTGGAGCTTTGGAG	2141

; CURRENT FILING DATE: 2003-02-10

; NUMBER OF SEQ ID NOS: 99

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 3

; LENGTH: 5075

; TYPE: DNA

; ORGANISM: Porcine parvovirus

US-10-361-004-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;

Best Local Similarity 70.4%; Pred. No. 2.1e-237;

Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

QY 4 GCTGGAAATGCTTACTCTGATCAAGTTTGGGAGCAACCAACTGCTTAAAGGAAAAAGT 63  
DB 298 GCGGGAACACATTACTTCGGAAGAGTACTAAAGACTCAACTGCTTCAGATATGCT 357  
QY 64 AACCGGAAGTCTTCTCAATTTGTTTTAAAAATGAAATGTTCAACTGAATGGAAGAAAT 123  
DB 358 CAAAAGAGCAATCTCTTATGATTTAAACACAAAAAGTCAATCTAAATGGAAGAA 417  
QY 124 ATCGGATGGAATAGTTACAAAAAGAGTCGAGGAGGAGCTGAATCTTTACACGA 183  
DB 418 ATTGCTTGAATAACTACAAAGATACACAGATGCGGAATGATAAACCTACAAAGA 477  
QY 184 GGAGCGAAACTACTTTGGGACCAAAAGCGAGGACATGGAATGGGAAACCAACAGTGGATGAA 243  
DB 478 GGAGCGAAACATCATGCGGACCGGCAACAGACATGGATGGGAATCAGAAATCGACAGC 537  
QY 244 ATGACCAAAAAGCAATATTCATTTTGGTAAATGTTTAAATGTTTAAATGTTTAAATGTT 303  
DB 538 CTCACAAAACGGCAAGTACTGATTTTGGTCTCTGTTTAAAAATGTTCTTTTGAAGGT 597  
QY 304 CTTAACACAAAGATATATTTCTGCTGATGTTAATGTTGTTGTCACACATGAATGGGA 363  
DB 598 ATATTGCAAAAGAACCTTAAGTCAAGTGACTGCTACTGTTTCATACAGATGAAATGTT 657  
QY 364 AAAGACCAAGGCTGGCACTGCCATGTAATAATTGGAGGAAAGGACTTTTAGTCAAGCTCAA 423  
DB 658 CAAGATACTGGCTATCACTGCCATGTAATAATTGGAGTGGAAAGGCTTACACAAAGCAATG 717  
QY 424 GGAATATGGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTTTAAACGCTGT 483  
DB 718 GGAATATGGTTCAGAAAAACAAATTAACAAATTTATGGAGTAGATGTTTAAATGCAATGC 777  
QY 484 AATGTCNACTACACAGCTGAAGAAATTAACATAAGAGAAATAGCAGAGCAATAGAG 543  
DB 778 AAAGTACCTCTAACACCAAGTTGAAGAATAAATTAAGGAATTACAGAGATGTTGAG 837  
QY 544 TGGGTTACTCTACTTATAGCATAAGCAACCAACCAAAAGACTATATACCAAGTGTGT 603  
DB 838 TGGGTATCGTACTAACCTACACTCACAAACAAACTAAAAACCAATATACAAAAATGACT 897  
QY 604 CTTTTTGGAAACATGATGTTTACTATATTTTAACTATAAAGAAATAAGCACTAGTCCA 663  
DB 898 CAITTTGGAAATATGATTTGTTTACTTCTTAAATAAATAAAGAAAGCAACT----- 951  
QY 664 CCAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGGAAACTAACTTTTAAAAA 723  
DB 952 GAAAGAGAGCATGATATATCTCAGCTCAGATCTGGCTTCATGACAAATTTCTTAAAA 1011  
QY 724 GAAAGCGAGCGCCATCTTAGTGAGCAACTATACATGATGACATGCGGCGCAGAAACGGTT 783  
DB 1012 GAAGCGGAGAGACACTTTAGTCAGTCACCTATTTTACTGAAGCAAAATAAACCTGAACTGTG 1071  
QY 784 GAAACCAAGTAACCTACCTGGCAGGAACTAAGCGGCGCAGAAATCAAACTAAAAAGAA 843  
DB 1072 GAAACCAAGTACTACAGCTCAGGAAGCCAAAAGAGGCGAGAAATACAAACAAAAAGAA 1131  
QY 844 GTTTCTATTAATACTACACTTAAAGAGCTGTTGATAAAAAGAGTAACCTCACCAGAGGAC 903  
DB 1132 GTAAGCATAAATGCACAAATAGAGACTTGGTTTAATAAAGATGTTACTAGCATAGAAGAC 1191

QY 904 TGGATGATGATGCGCCAGACAGTTACATTGAATGATGGCTCAACAGGTGGAGAAAC 963  
DB 1192 TGGATGATGACAGATCCAGACAGTTATATAGAAATGATGGCTCAACCGGAGGAAAT 1251  
QY 964 CTGCTGAAAAATAGCTAGAGATTTGTACACTTAATCTAGCCAGAACCAAAAAACAGATTT 1023  
DB 1252 TTAATCAAAATACACTAGAAATTAACAATCTTACTCTAGCAGAGAACAAAAACAGCATAT 1311  
QY 1024 GACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAAACCAACTTTTCACTGCCTGACACA 1083  
DB 1312 GACTTAATACTTTGAAAAAGCAAAACCAAGCATGTACCAACATTTAATATTAGCAATACA 1371  
QY 1084 AGAGCTGCGAATTTTGTCTTTTCACTGGCTGGAACTATGTTAAAGTTTGGCATGCTTATT 1143  
DB 1372 AGAACATGTTAAATATTTCAAGCATGCAAACTTGGAACTACATTAAGTCTGCGCATGCTATA 1431  
QY 1144 TGCTGTGTTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTTTATTTTATTTTATGGACCGC 1203  
DB 1432 ACTTGTGTACTTAACAGACAGGAGGAAAGAAATACAAATTTCTATTTCAITGGGCCAGCA 1491  
QY 1204 AGCAGCGCAAAATCTATTATTGCACAAGCCATAGCACAAGCAGTTGGCAATGTTGGTTGC 1263  
DB 1492 TCAACAGGAAAAAGTATAATTTGCTCAACACATTTGCAAACTTAGTTGGTAATTTGGTTGC 1551  
QY 1264 TATTAATGCGCAATGTAACCTTTTCCATTTAATGACTGTACCAACAAAGAACTTGCATTTGG 1323  
DB 1552 TACAATGCGCCAAATGTGAATTTTCCATTTAATGACTGTACAAATTAATAAACTTTAATATGG 1611  
QY 1324 GTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGT 1383  
DB 1612 ATTGAGAGAGCAGGAACTTCTCTAAACCAAGTAAACCAATTTCAAGCCATATGTTTCAGGT 1671  
QY 1384 CAAACTATTTCGATTTGATCAAAAAAGGAAAGCGAGCAAAAGAGTTGAACCAACACCAAGTC 1443  
DB 1672 CAAACAAATTAGAAATTTGACCAAAAAAGTAAAGAGCAAAACAAATTTGAACCAACTCTGTA 1731  
QY 1444 ATCATGACCACAATAGAGCAATTTACAGTGTGAGATAGCTGCGAAGAGAAAGACCAAGAA 1503  
DB 1732 ATAATGACTACAAATGAAGACATTAACAGTTAGATAGGATGCGAGGAAAGACCAAGAA 1791  
QY 1504 CACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTTAAACATATACCTTGCCTGGT 1563  
DB 1792 CATACACAACTAATAAGAGACAGAAATGTTAAACATAAACCTTAAACCAAGAAATGCGCAGT 1851  
QY 1564 GACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGCTTGGTTGTTAAAGAAAT 1623  
DB 1852 GATTTTGGACTTTTAGAAGAAAACTGAATGGCCACTTAATATGCTTGGTTGGTTGAAGAAA 1911  
QY 1624 GGTACCACATCTPACCATGCGAAGCTACTGTGCTAAATGGGCAAAAGTTTCTGTAITGGTCA 1683  
DB 1912 GGTACCAGCAACAAATGGCTAGCTATATGCAATGTTGGGAAATGTACCTGATTTGTTCA 1971  
QY 1684 GAAACTGGGCGGAGCCAAAGGTGCAACTCTCTATAATTTTACTAGTTTGGCAGCTGCA 1743  
DB 1972 GAAATGGGAGGAGCCAAAAATGCAACCCCAATAAATAACACCA---ACAGACTCTCAG 2028  
QY 1744 CAATTCACGACACCAAGAAAGTACGCTCTCAGCCAGAACTATGCTAACTACCTCCACTTGCA 1803  
DB 2029 ATTTCCACATAGTGAATACTTCGCGCAGCGGCAACAACTACGCGAGAACTCCAATACAG 2088  
QY 1804 TCGGATCTCGA---GGACCTGGCTTTTAGAGCTTTGGAGCACACCAAAATACTCC 1853  
DB 2089 GAGGACCTGGATTTAGCTTTAGCCTTTGGAGCCGTGGAGCGGAGCCCAACACACC 2141

## RESULT 4

US-10-466-894-1276

; Sequence 1276, Application US/10466894

; Publication No. US20040241671A1

; GENERAL INFORMATION:

; APPLICANT: Teletman, Adam

; APPLICANT: Amson, Robert

; APPLICANT: Tuijnder, Marius

APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1276  
; LENGTH: 374  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-1276

Query Match 15.2%; Score 307.4; DB 18; Length 374;  
Best Local Similarity 88.2%; Pred. No. 1.6e-70; Indels 0; Gaps 0;  
Matches 329; Conservative 0; Mismatches 44;

QY 174 TTTTACACGAGGCGGAACTACTTTGGGACCAAGCGAGCATGGAATGGGAACAC 233  
Db 1 TTTTACACGAGGCGGAGACCACTTTGGGACCAAGCGAGCATGGAATGGGAGCGC 60  
QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGTAAATAATGTTT 293  
Db 61 AGTGGATGATGACCAAAAGCAAGTATTTATTTTGGTAAATAATGTTT 120  
QY 294 ATTGGAAGTCTTAACACAAAGATATTTCTGCTGATGTTAAATGTTGCAACA 353  
Db 121 GTTTGAAGTCTGACACAAAGACATAGCTCTAGTAATGTTTCTGTCGACGA 180  
QY 354 TGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGCACTTTAG 413  
Db 181 TGAATGGGAAAGACCAAGGCTGGCACTGTCTGATTTGGAGGCAAGCACTTTAG 240  
QY 414 TCAAGCTCAAGGAATGTTGGAGAAAGCACTAAATGTTTACTGGAGAGATGTTGTT 473  
Db 241 TCAACCTCAAGGAATGTTGGAGAAAGCACTAAATGTTTACTGGAGTANATGTTG 300  
QY 474 ACACGCTGATGTCACACTACACCAAGCTGAAAGTAACTAAAGAGAAATAGCAGA 533  
Db 301 GACTGCTGNAATGTTNCAACTAACCAAGCTGAAAGAAATTAACCTGNNAGAAATAGCAGA 360  
QY 534 AGACAAATGAGTGG 546  
Db 361 GGACAGTGANNG 373

RESULT 5  
US-10-466-894-430  
; Sequence 430, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Teلمان, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijinder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 430  
; LENGTH: 421  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: 276, 405, 417, 419, 421  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-430

Query Match 14.2%; Score 285.8; DB 18; Length 421;  
Best Local Similarity 87.6%; Pred. No. 8.9e-65;  
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTTGGGACCAAGCGAGCATGGAATGGGAACCA 232  
Db 59 CTTTACACGAGGCGGAGACCACTTTGGGACCAAGCGAGCATGGAATGGGAGCG 118  
QY 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCATTTTGGTAAATAATGTT 292  
Db 119 CAGTGGATGATGACCAAAAGCAAGTATTTATTTTGGTAAATAATGTT 178  
QY 293 TATTTGAAGTCTTAACACAAAGATATTTCTGCTGATGTTAAATGTTTGTCAAC 352  
Db 179 TGTTTGAAGTCTGACACAAAGACATAGCTCCTAGTAATGTTTCTGTTGTCGACG 238  
QY 353 ATGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGCACTTTA 412  
Db 239 ATGAATGGGAAAGACCAAGGCTGGCACTGTCTGATTTGGAGGCAAGCACTTTA 298  
QY 413 GTCAGCTCAAGGGAATGTTGGAGAAAGCACTAAATGTTTACTGAGCAGATGTTGG 472  
Db 299 GTCACCTCAAGAAATGTTGGAGAAAGCACTAAATGTTTACTGAGTAGATGTTGG 358  
QY 473 TAACAGCTGTAATGTGCAACTAACCAAGCTGAAAGAAATTAACCTAAGAGAAAT 527  
Db 359 TGACTGCTGTAATGTTCAACTAACCAAGCTGAAAGAAATTAACCTAAGAGAAAT 413

RESULT 6  
US-10-466-894-375  
; Sequence 375, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Teلمان, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijinder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 375  
; LENGTH: 423  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: 422  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-375

Query Match 14.1%; Score 284.8; DB 18; Length 423;  
Best Local Similarity 89.2%; Pred. No. 1.6e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232  
DB 71 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGGAGAGCG 130

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATCTTTTGTGTTAAAGATGTT 292  
DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATCTTTTGTGTTAAAGATGTT 190

QY 293 TATTTTGAAGTCTTAACACAAAGATATATTTTCTGGTGTGATCTTAATTTGGTTGTGCAAC 352  
DB 191 TGTTTGAAGTCTCAGCACAAAGACATAGCTCTTAAGTATTTACTTGTGTCGTCGACG 250

QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGGCATGTACTTAATGGAGAAAGGACTTTA 412  
DB 251 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTCTGATTTGGAGGCAAGGACTTTA 310

QY 413 GTCAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
DB 311 GTCAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTGGAGTATGTTGGTGG 370

QY 473 TAACAGCTGTAAATGTGCAACTAAACACAGCTGAAAGAAATTTAA 516  
DB 371 TGACTGCCCTGTAAATGTTCAACTAAACACAGCTGAAAGAAATTTAA 414

RESULT 7  
US-10-466-894-370  
; Sequence 370, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telesman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 370  
; LENGTH: 464  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 414, 435, 459, 462, 464  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-370

Query Match 14.1%; Score 284.8; DB 18; Length 464;  
Best Local Similarity 89.2%; Pred. No. 1.7e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232  
DB 55 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGGAGAGCG 114

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTTTGTGATCTTTTGTGTTAAAGATGTT 292  
DB 115 CAGTGGATGACATGACCAAAAGCAAGTATTTTGTGATCTTTTGTGTTAAAGATGTT 174

QY 293 TATTTGAAGTCTTAAACACAAAGATATATTTTCTGGTGTGATTTAATTTGGTTTGTGCAAC 352  
DB 175 TGTTTGAAGTCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTTACTTTGTTCTGTGCAGC 234

QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGGCATGTACTTAATTTGGAGAAAGGACTTTA 412  
DB 235 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTCTGATTTGGAGGCAAGGACTTTA 294

QY 413 GTCAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
DB 295 GTCAGCTCAAGGAAATGGTGAGAGGCAAGTAAATGTTGTTCTGGAGTATGTTGGTGG 354

QY 473 TAACAGCTGTAAATGTGCAACTAAACACAGCTGAAAGAAATTTAA 516  
DB 355 TGACTGCCCTGTAAATGTTCAACTAAACACAGCTGAAAGAAATTTAA 398

RESULT 8  
US-10-466-894-367  
; Sequence 367, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telesman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 367  
; LENGTH: 473  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 10, 431, 447, 467, 468, 473  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-367

Query Match 14.1%; Score 284.8; DB 18; Length 473;  
Best Local Similarity 89.2%; Pred. No. 1.8e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232  
DB 72 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGGAGAGCG 131

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATCTTTTGTGTTAAAGATGTT 292  
DB 132 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATCTTTTGTGTTAAAGATGTT 191

QY 293 TATTTTGAAGTCTTAACACAAAGATATATTTTCTGGTGTGATTTAATTTGGTTTGTGCAAC 352  
DB 192 TGTTTGAAGTCTCAGCACAAAGACATAGCTCTCTAGTAATGTTTACTTGGTTGTGTCAGC 251

QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGGCATGTACTTAATTTGGAGAAAGGACTTTA 412  
DB 252 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTCTGATTTGGAGCAAGGACTTTA 311

QY 413 GTCAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
DB 312 GTCAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTGGAGTATGTTGGTGG 371

QY 473 TAACAGCTGTAAATGTGCAACTAAACACAGCTGAAAGAAATTTAA 516



```

; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 418
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-438

Query Match
Best Local Similarity 14.1%; Score 284.2; DB 18; Length 420;
Matches 310; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 173 CTTTCAACGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACCA 232
Db 64 CTTTCAACGAGGAGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAAGCG 123
QY 233 CAGTGGATCAATGACCAAAAGCAAGTATTCATTTTGTGATTTTGTGTTTAAAAAATGTT 292
Db 124 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTTGTGTTTAAAAAATGTT 183
QY 293 TATTTGAAGTGCTTAACACAAAGATATATTTCTCGGTGATGTTAATTTGGTTTGGCAAC 352
Db 184 TGTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTTACTTTGGTTTGTGCAGC 243
QY 353 ATGAATGGGAAAGACCAAGCTGGCACTGCCATGCTACTAATTTGAGGAAAGGACTTTA 412
Db 244 ATGAATGGGAAAGACCAAGCTGGCACTGTCATGCTGATTTGAGGCAAGGACTTTA 303
QY 413 GTCAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTTGGACAGATGTTGG 472
Db 304 GTCAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTTGGAGTAGATGTTGG 363
QY 473 TAACAGCTGTAAATGTGCAACTAACACCAAGCTGAAGAATTAACCTAAAGAA 525
Db 364 TGACTGCTGTAAATGTGCAACTAACACCAAGCTGAAGAATTAACCTAAAGAA 416

RESULT 12
US-10-466-894-388
; Sequence 388, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 388
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-388

Query Match
Best Local Similarity 14.1%; Score 283.8; DB 18; Length 343;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGAGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAAGCGC 60
QY 234 AGTGGATGAATGACCAAAAGCAAGTATTTCTTTGATTTCTTTGTTTAAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTCTTTGTTTAAAGAGTTT 120
QY 294 ATTTGAAGTGCTTAACACAAAGATATATTTCTCGGTGATGTTAAATTTGGTTTCTGCAACA 353
Db 121 GTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTTACTTTGTTCTGTCAGCA 180
QY 354 TGAATGGGAAAGACCAAGCTGGCACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGGACCAAGCTGGCACTGTCATGCTGATTTGGAGGCAAGCACTTTAG 240
QY 414 TCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATTTTACTGAGCAGATGTTGGT 473
Db 241 TCAACCTCAAGGAAATGTTGAGAGGCACTAAATTTGTTACTGAGTAGATGTTGGT 300
QY 474 AACAGCTGTAAATGTGCAACTAACACCAAGCTGAAGAATTA 516
```

```

QY 234 AGTGGATGAATGACCAAAAGCAAGTATTTCTTTGATTTCTTTGTTTAAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTCTTTGTTTAAAGAGTTT 120
QY 294 ATTTGAAGTGCTTAACACAAAGATATATTTCTCGGTGATGTTAAATTTGGTTTCTGCAACA 353
Db 121 GTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTTACTTTGTTCTGTCAGCA 180
QY 354 TGAATGGGAAAGACCAAGCTGGCACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGGACCAAGCTGGCACTGTCATGCTGATTTGGAGGCAAGCACTTTAG 240
QY 414 TCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATTTTACTTGGAGCAGATGTTGGT 473
Db 241 TCAACCTCAAGGAAATGTTGAGAGGCACTAAATTTGTTACTTGGAGTAGATGTTGGT 300
QY 474 AACAGCTGTAAATGTGCAACTAACACCAAGCTGAAGAATTA 516
Db 301 GACTGCTGTAAATGTGCAACTAACACCAAGCTGAAGAATTA 343

RESULT 13
US-10-466-894-390
; Sequence 390, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-390

Query Match
Best Local Similarity 14.1%; Score 283.8; DB 18; Length 343;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGAGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAAGCGC 60
QY 234 AGTGGATGAATGACCAAAAGCAAGTATTTCTTTGATTTCTTTGTTTAAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTCTTTGTTTAAAGAGTTT 120
QY 294 ATTTGAAGTGCTTAACACAAAGATATATTTCTCGGTGATGTTAAATTTGGTTTCTGCAACA 353
Db 121 GTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTTACTTTGTTCTGTCAGCA 180
QY 354 TGAATGGGAAAGACCAAGCTGGCACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGGACCAAGCTGGCACTGTCATGCTGATTTGGAGGCAAGCACTTTAG 240
QY 414 TCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATTTTACTGAGCAGATGTTGGT 473
Db 241 TCAACCTCAAGGAAATGTTGAGAGGCACTAAATTTGTTACTGAGTAGATGTTGGT 300
QY 474 AACAGCTGTAAATGTGCAACTAACACCAAGCTGAAGAATTA 516
```

Db 301 GACTGCTGTAAATGTTCAACTAACACGCTGAAAAA 343

RESULT 14

US-10-466-894-433  
; Sequence 433, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 433  
; LENGTH: 403  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 218  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-433

Query Match 14.1%; Score 283.8; DB 18; Length 403;  
Best Local Similarity 89.0%; Pred. No. 2.9e-64;  
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Qy 173 CTTTACACGAGGAGCGGAAACTTCTTGGACCAAGCGAGGACATGGAAATGGGAAACCA 232  
Db 56 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAAATGGGAGCG 115  
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAAATGTT 292  
Db 116 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAAATGTT 175  
Qy 293 TATTTGAAGTGTCTTACACAAAGAAATATATTTCTGCTGATGTTAATTTGGTTGTCAC 352  
Db 176 TGTTCGAAGTGTCTTACACAAAGAAATATATTTCTGCTGATGTTAATTTGGTTGTCAC 235  
Qy 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCTTAAATTTGGAGGAAAGGACTTTTA 412  
Db 236 ATGAATGGGAAAGAACCAAGGCTGGCACTGCTTAAATTTGGAGGAAAGGACTTTTA 295  
Qy 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCACTAAATTTTACTGGAGCAGATGTTGG 472  
Db 296 GTCAAGCTCAAGGAAATGGTGGAGAGGCACTAAATTTTACTGGAGCAGATGTTGG 355  
Qy 473 TAACAGCTGTAAATGTGCAACTAACACGCTGAAAGAAATTTAA 516  
Db 356 TGACTGCTGTAAATGTGCAACTAACACGCTGAAAGAAATTTAA 399

RESULT 15

US-10-466-894-369  
; Sequence 369, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 369  
; LENGTH: 497  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 448-475  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-369  
Query Match 14.0%; Score 283.2; DB 18; Length 497;  
Best Local Similarity 89.0%; Pred. No. 4.8e-64;  
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Qy 173 CTTTACACGAGGAGCGGAAACTTCTTGGACCAAGCGAGGACATGGAAATGGGAAACCA 232  
Db 68 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAAATGGGAGCG 127  
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAAATGTT 292  
Db 128 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGAATCTTTTGGTTAAAAATGTT 187  
Qy 293 TATTTGAAGTGTCTTACACAAAGAAATATATTTCTGCTGATGTTAATTTGGTTGTCAC 352  
Db 188 TGTTCGAAGTGTCTTACACAAAGAAATATATTTCTGCTGATGTTAATTTGGTTGTCAC 247  
Qy 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCTTAAATTTGGAGGAAAGGACTTTTA 412  
Db 248 ATGAATGGGAAAGAACCAAGGCTGGCACTGCTTAAATTTGGAGGAAAGGACTTTTA 307  
Qy 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCACTAAATTTTACTGGAGCAGATGTTGG 472  
Db 308 GTCAAGCTCAAGGAAATGGTGGAGAGGCACTAAATTTTACTGGAGCAGATGTTGG 367  
Qy 473 TAACAGCTGTAAATGTGCAACTAACACGCTGAAAGAAATTTAA 516  
Db 368 TGACTGCTGTAAATGTGCAACTAACACGCTGAAAGAAATTTAA 411

Search completed: January 23, 2005, 14:35:11  
Job time : 1074 secs

11

12

13

14

15



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:58 ; Search time 6354.75 Seconds  
(without alignments)  
11577.453 Million cell updates/sec

Title: us-10-069-056-8

Perfect score: 2019

Sequence: 1 atggctggaatcttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	61.6	3.1	1101	9	CNS0039G
C 2	59.8	3.0	997	9	CNS005TE
C 3	50	2.5	1134	8	CC245800 CH261-4F2
C 4	49.8	2.5	1101	9	CNS0006J
C 5	49.8	2.5	1101	9	CNS0182P
C 6	49.6	2.5	1101	9	CNS017KX
C 7	49.4	2.4	886	9	CNS075DF
C 8	49.2	2.4	913	9	CG772152
C 9	49.2	2.4	930	9	CG770435
C 10	48.4	2.4	1063	9	CG770435 TC852.2_A
C 11	48.4	2.4	1241	9	AG436064 T3 end of
C 12	47.8	2.4	987	9	AG436064 T3 end of
C 13	47.2	2.3	1025	9	CNS00418
C 14	47.2	2.3	1811	9	CNS014U2
C 15	46.6	2.3	933	9	CG753732 P048-4-G0
C 16	46.4	2.3	1080	9	AL059220 Drosophila
C 17	46.2	2.3	1758	9	AL059220 Drosophila
C 18	46	2.3	769	9	CL509408 SAIL_811
C 19	45.8	2.3	963	9	AL100817 Drosophila
C 20	45.8	2.3	1066	6	CNS00A4L
C 21	45.8	2.3	1204	9	CD386086 AGENCOURT
C 22	45.6	2.3	1200	9	AL106628 Drosophila
C 23	45.4	2.2	975	6	AL106647 Drosophila
C 24	45.2	2.2	644	8	CD050911 AGENCOURT
					BH463121 BOHEK14TR

C 25	45.2	2.2	979	8	AZ538989
C 26	45	2.2	768	9	ENTPH71TF
C 27	45	2.2	894	9	AL186863 Tetracodon
C 28	45	2.2	1184	8	AL109126 Drosophila
C 29	44.8	2.2	368	2	B13117 T9K3-Sp6.2
C 30	44.8	2.2	414	4	AW767845 da80h01.x
C 31	44.8	2.2	435	8	BJ640957
C 32	44.8	2.2	581	9	BZ665713 KJ640957
C 33	44.6	2.2	298	4	BZ665713 KJ640957
C 34	44.6	2.2	363	5	B1815258 PFESToab0
C 35	44.6	2.2	870	8	BQ451492 PFESToab0
C 36	44.4	2.2	706	9	AZ693037 ENTHP79TF
C 37	44.4	2.2	878	9	AG178839 Pan trogl
C 38	44.4	2.2	938	9	AL108993 Drosophila
C 39	44.4	2.2	1201	5	AL085906 Drosophila
C 40	44.2	2.2	588	1	AX416077 BX416077
C 41	44.2	2.2	694	1	AJ773389 AJ773389
C 42	44.2	2.2	1500	9	AJ778527 AJ778527
C 43	44	2.2	871	5	AG334092 Mus muscu
C 44	44	2.2	922	9	BU221434 603107334
C 45	44	2.2	1101	9	CG766473 TCB41.1.H
					AL061823 Drosophila

#### ALIGNMENTS

RESULT 1  
CNS0039G/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
AL063921

VERSION  
AL063921.1 GI:4941778

KEYWORDS  
GSS.

SOURCE  
Drosophila melanogaster (fruit fly)

ORGANISM  
Drosophila melanogaster

REFERENCE  
1 (bases 1 to 1101)

AUTHORS  
Genoscope.

TITLE  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

JOURNAL  
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT  
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeogawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, N.Y. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
Location/Qualifiers  
1..1101

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone\_lib="BACR08K10"

/note="end : TET3"

Query Match 3.1%; Score 61.6; DB 9; Length 1101;

Best Local Similarity 16.2%; Pred. No. 0.00016;







**KEYWORDS**  
Drosophila melanogaster (fruit fly)

**SOURCE ORGANISM**  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.

**REFERENCE AUTHORS**  
**TITLE**  
**JOURNAL**  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

**COMMENT**  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

**FEATURES**

source	Location/Qualifiers
	1..1101
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACN37F10"
	/clone_lib="DrosBAC"
	/plasmid="pBelobAC11"
	/note="end : Sp6"

**ORIGIN**

Query Match	2.5%	Score 49.6;	DB 9;	Length 1101;
Best Local Similarity	21.8%;	Pred. No. 0.21;		
Matches	95;	Conservative 143;	Mismatches 198;	Indels 0; Gaps 0;

```

QY      260 TATTCAATTTTGGTCTTCCTTGTTGTTTAAATAAATGTCTTAATTTGAAGTCTTAACACAAGAATA 319
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1077 TWKTADKKTRWDTKDGCGGGGKGKGGDKRKAADTKRDWRGAATRRAARWATAATTATA 1018
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      320 TATTTCCTGGTGATGTTAAATGGTTTCTGCACCAATGAATCGGGAAGAACCAAGGCTGGC 379
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1017 DDTKRTTGTGTRTWKWKWWARRAKRWDRDWTDWDKWDTKWDKDDRTKRWAATRGDGR 958
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      380 ACTGCCATGTACTAAATTTGGAGGAAGACATTTTAGTCAAGCTCAAGGGAATGGTGAGAA 439
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      957 AGRKRRDDRGRGRERRRRRTKFWRRRAAWWTAAWTTWNGTTKGAWWTKRKCKG 898
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      440 GGCAACTAAATGTTTACTCGAGCAGATGGTTGGTAAACAGCGCTGTAATGTGCAACTAACAC 499
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      897 GA AAAAADWGRTDDWRWDKRAAAAAGKDGKGKARTWGA AAAAAA WDTTKTATKKRKT 838
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      500 CAGCTGAACAAATTAACACTAAGAAATACGACAGACAATGAGTGGTGTACTCTACTTA 559
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      837 KTGRAMWAAGTRITWDAATATKTKRAAAAAGRRAAAAWAKDRGTKRKGKRGTTDDGKG 778
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      560 CTTATAGCATATAAGCAAACCAAAAAGACATATACCAAGTGTGTCTTTTTGGAAAACATGA 619
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      777 TKWKTRITWTKTAAARBARABAAAKGWTKTKTDGRRARAANDANTTYTRKGATRDGAA 718
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      620 TTGCTTACTATTTTTTAACTAA AAAGAAAATAAGCACTAGTCCACCAAGAGACGGAGCT 679
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      717 AAAATTGGAAAAAATGGKTTDAGAAAAATAAAWRRTTWTDWAAAAAAGAAAWAKAAAWD 658
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      680 ATTTCTTACAGTGA 695
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      657 AAADA KKKGRRAKRA 642
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

**RESULT 7**  
CNS075DF LOCUS  
CNS075DF DEFINITION  
clone BA0AB034H01 of library BA0AB from strain CLIB 210 of Kluyveromyces lactis, genomic survey sequence. linear GSS 07-JUL-2001

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AL429961	AL429961.1	GI:12213155	GSS.	
REFERENCE	AUTHORS			
				Kluyveromyces lactis
				Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
				1 (bases 1 to 886)
				Souciot J.L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brottier, P., Cagaregola, S., de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuveglise, C., Olier-Kalogeropoulos, O., Potier, S., Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P. and Weissenbach, J.
				Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
				FEBS Lett. 487 (1), 3-12 (2000)
				20584711
				11152876
				2 (bases 1 to 886)
				Bolotin-Fukuhara, M., Toffano-Nioche, C., Artiguenave, F., Duchateau-Nguyen, G., Lemaire, M., Marmeisse, R., Montrocher, R., Robert, C., Ternier, M., Wincker, P. and Wesolowski-Louvel, M.
				Genomic exploration of the hemiascomycetous yeasts: 11.
				Kluyveromyces lactis
				20584721
				11152885
				3 (bases 1 to 886)
				Genoscope.
				Direct Submission
				Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
				This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers			
	1..886			
	/organism="Kluyveromyces lactis"			
	/mol_type="genomic DNA"			
	/strain="CLIB 210"			
	/variety="lactis"			
	/db_xref="taxon:28985"			
	/clone="BA0AB034H01"			
	/clone_lib="BA0AB"			
	complement(<619..>886).			
	/note="similar to Saccharomyces cerevisiae ORF YNR001c [CIT1 ; citrate (si)-synthase, mitochondrial]"			
	/evidence=not_experimental			
misc_feature				
	2.4%; Score 49.4; DB 9; Length 886;			
	Best Local Similarity 35.5%; Pred. No. 0.22;			
	Matches 196; Conservative 64; Mismatches 291; Indels 1; Gaps 1;			
ORIGIN				
	491 AACTACACAGCTGGAAGAAATTAACACTAAGAGAAATAGCAGAGCAATGCTGGTTA 550			
	Db 34 AAAAAACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 93			
	551 CTCTACTTACTTATAGCATAAGCAACCAAAAGACCTATACCAAGTGTGTTCTTTG 610			
	Db 94 AAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 153			
	611 GAAACATGATTCCTACTATTTTTTAACATAAAGAAAAATAAGCACTAGTCCACCAAGAG 670			

```
Db 154 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 213
QY 671 ACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAAAGAGGCG 730
Db 214 AAAAAATRTTGAGAPAAATRWITGRTGTGTAWRWWRWRWRTTATRWATRAAGWR 273
QY 731 AGCGCATCTAGTGAGCAAACTATACACTGTATGACATGCGGCGAGAAACGGTTGAAACCA 790
Db 274 AAWRAAAATGWRAGGAGGAGGAAGAAGARTATATAAARAAGARAAPAAATWAAAWAR 333
QY 791 CA-GTAACCACTGCGGAGAAACTAAGCGCGGAGAAATTCAACTAAAGAAAGTTCT 849
Db 334 TAGGTWAAAAAGTATMARWAWAWAWAAGAAAAAAMAAAAAAMAAAAAAMAAAA 393
QY 850 ATTAAACCTACACTTAAAGAGCTGGTGCAATAAAGAGTAACCTCACGAGGACTGGATG 909
Db 394 CRGMAAAMAAAMAAACCGTAGAARGAAAMAAAMAAAMAAAMAAAMAAAMAA 453
QY 910 ATGATGACGCCAGACAGTTTACATTGAAATGATGGCTCAACCGAGTGGAGAAAACCTGCTG 969
Db 454 AARGACTTWRGMAWAWAAGATAGAAAAMCAACGMAAAAMGAARAAMAAAMAAAGMAAR 513
QY 970 AAAAAACCTAGAGATTTGTACACTTAACCTTAGCCAGAACCAAAACACACATTTGACTTA 1029
Db 514 AMACMAAAACAMAAAMAAACAMMAAAAAACAGAAAAAAMAAAAAAMAAACAAACCR 573
QY 1030 ATTTTAAAAA 1041
Db 574 AAGTAAAAA 585

RESULT 8
CG772152 913 bp DNA linear GSS 29-OCT-2003
LOCUS TcB48.2_H01_T7 Tribolium BAC library Tribolium castaneum genomic,
DEFINITION genomic survey sequence.
ACCESSION CG772152 GI:38025852
VERSION
KEYWORDS
SOURCE
ORGANISM
Tribolium castaneum (red flour beetle)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
REFERENCE 1 (bases 1 to 913)
AUTHORS Savard, J. and Tautz, D.
TITLE Tribolium castaneum BAC-ends sequencing project
JOURNAL Unpublished (2003)
COMMENT Abteiling fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..913
/organism="Tribolium castaneum"
/mol_type="genomic DNA"
/strain="GA-2"
/db_xref="taxon:7070"
/clone_lib="Tribolium BAC library"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match 2.4%; Score 49.2; DB 9; Length 913;
Best Local Similarity 58.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 63;

QY 23 ATGAAGTTTGGGAGCAACCACTGGTTAAAGAAAAAGTAACCCAGGAGTGTCTCAT 82
Db 288 ATGAATTTTACACAGAGCAATTAATACATCTTTAAATTCATTTGGAGTATTCTGAA 347
QY 83 TTGTTTTTAAAAATGAAATGTTCACTGAATGGAAGATATCGGATCGAATAGTTACA 142
Db 348 CTATTTTATAACAAAAATCTTCGAAAGACAGAAAAATAAATATCGACTAACTAAA 407
QY 143 AAAAGAGCTGCAGGAGGACGAGCTGAAT 172
Db 408 AAATAACTATCGACAACTCTGTACTAAAT 437

RESULT 10
CNS07A2Y 1063 bp DNA linear GSS 08-JUL-2001
LOCUS T3 end of clone BC0AA001B01 of library BC0AA from strain CBS 767 of
DEFINITION Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL436064
VERSION AL436064.1 GI:12219477
KEYWORDS
SOURCE Debaryomyces hansenii (anamorph: Candida famata)
```

```
Db 296 ATGAATTTTACACAGAGCAATTAATTAATCTTTAAATGATTTGGAGTATTCTGAA 355
QY 83 TTGTTTTTAAAAATGAAATGTTCACTGAATGGAAGATATCGGATCGAATAGTTACA 142
Db 356 CTATTTTATAACAAAAATCTTCGAAAGACAGAAAAATAAATATCGACTAACTAAA 415
QY 143 AAAAGAGCTGCAGGAGGACGAGCTGAAT 172
Db 416 AAATAACTATCGACAACTCTGTACTAAAT 445

RESULT 9
CG770435 930 bp DNA linear GSS 29-OCT-2003
LOCUS TcB52.2_A06_Sp6 Tribolium BAC library Tribolium castaneum genomic,
DEFINITION genomic survey sequence.
ACCESSION CG770435
VERSION
KEYWORDS
SOURCE
ORGANISM
Tribolium castaneum (red flour beetle)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
REFERENCE 1 (bases 1 to 930)
AUTHORS Savard, J. and Tautz, D.
TITLE Tribolium castaneum BAC-ends sequencing project
JOURNAL Unpublished (2003)
COMMENT Abteiling fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..930
/organism="Tribolium castaneum"
/mol_type="genomic DNA"
/strain="GA-2"
/db_xref="taxon:7070"
/clone_lib="Tribolium BAC library"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match 2.4%; Score 49.2; DB 9; Length 930;
Best Local Similarity 58.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 63;

QY 23 ATGAAGTTTGGGAGCAACCACTGGTTAAAGAAAAAGTAACCCAGGAGTGTCTCAT 82
Db 288 ATGAATTTTACACAGAGCAATTAATTAATCTTTAAATTCATTTGGAGTATTCTGAA 347
QY 83 TTGTTTTTAAAAATGAAATGTTCACTGAATGGAAGATATCGGATCGAATAGTTACA 142
Db 348 CTATTTTATAACAAAAATCTTCGAAAGACAGAAAAATAAATATCGACTAACTAAA 407
QY 143 AAAAGAGCTGCAGGAGGACGAGCTGAAT 172
Db 408 AAATAACTATCGACAACTCTGTACTAAAT 437

RESULT 10
CNS07A2Y 1063 bp DNA linear GSS 08-JUL-2001
LOCUS T3 end of clone BC0AA001B01 of library BC0AA from strain CBS 767 of
DEFINITION Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL436064
VERSION AL436064.1 GI:12219477
KEYWORDS
SOURCE Debaryomyces hansenii (anamorph: Candida famata)
```

```

ORGANISM  Debaromyces hansenii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Debaromyces.
REFERENCE 1 (bases 1 to 1063)
AUTHORS  Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
          Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
          de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
          Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
          Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
          Wincker,P. and Weissenbach,J.
TITLE  Genomic exploration of the hemiascomycetous yeasts: 1. A set of
        yeast species for molecular evolution studies
JOURNAL  FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE  20584711
PUBMED  11152876
REFERENCE 2 (bases 1 to 1063)
AUTHORS  Lepingle,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H.,
          Artiguenave,F., Wincker,P. and Gaillardin,C.
TITLE  Genomic exploration of the hemiascomycetous yeasts: 14.
        Debaromyces hansenii var. hansenii
JOURNAL  FEBS Lett. 487 (1), 82-86 (2000)
MEDLINE  20584724
PUBMED  11152889
REFERENCE 3 (bases 1 to 1063)
AUTHORS  Genoscope.
TITLE  Direct Submission
JOURNAL  Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
        2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
        seq@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT  This GSS is part of a random genomic sequencing program of thirteen
        yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
        exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
        Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
        lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
        angusta, Debaromyces hansenii var. hansenii, Pichia sorbitophila,
        Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
        5 kb were prepared and both extremities were sequenced. See
        keywords for description of this sequence and for the sequence of
        the other extremity of this insert.
FEATURES  Location/Qualifiers
          source            1..1063
                           /organism="Debaromyces hansenii"
                           /mol_type="genomic DNA"
                           /strain="CBS 767"
                           /variety="hansenii"
                           /db_xref="taxon:4959"
                           /clone="BC0AA001B01"
                           /clone_lib="BC0AA"
                           /notes="end : T3"
ORIGIN
Query Match      2.4%; Score 48.4; DB 9; Length 1063;
Best Local Similarity 38.2%; Pred. No. 0.42;
Matches 108; Conservative 38; Mismatches 137; Indels 0; Gaps 0;
QY 369 CCAAGCTGTCATGTCATGTAATAATGGAGAAAGGACTTTAGTCAAGCTCAAGCGAA 428
Db 661 CCCGSSSSSACCGGCTCGTCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
QY 429 ATGCTGCGAGAGGCACTAAATCTTTTCTGACAGATGTTGGTAAACAGCCTGTAATGT 488
Db 721 TAAATGGATATTTTRDDGGTTTGTATAAGGGGGRRTATRTDTKTCTTTGGGGGGGRRRA 780
QY 489 GCAACTAACACCACTGTAAGAAATTAACCTAAGAGAAATAGCAGAGCAATCAGTGGGT 548
Db 781 RRAPRRRGAAAGGRRAGGGGKGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 840
QY 549 TACTCTACTTACTTATAAGCATAGCAACCAAAAGAGACTATACCAAGTGTGTTCTTTT 608
Db 841 AAATTTTTTAAAAAGAAAAAGAAATATAAGAAAAAGAAATTTTTTWTATTTTATATTTT 900
QY 609 TGAACATGATGCTTACTATTTTAACTAAAGAAATA 651

```

```

Db 901 TTTTATATATATTTTAAATTTTATTTTWTATTAAGAAAAA 943

RESULT 11
AG448181
LOCUS   AG448181
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-332A17.T7, genomic survey
          sequence.
ACCESSION AG448181
VERSION   AG448181.1 GI:48091244
KEYWORDS  GSS.
SOURCE    Mus musculus molossinus
ORGANISM  Mus musculus molossinus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE    BAC end Sequences of Library MSMg01
JOURNAL  Unpublished
AUTHORS  2 (bases 1 to 1241)
TITLE    Direct Submission
JOURNAL  Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT  Clones are derived from the mouse BAC library MSMg01. For BAC
          library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
          Tsukuba Institute, Bio Resource Center, Research (RIKEN) 3-1-1
          The Institute of Physical and Chemical Research (RIKEN) 3-1-1
          Koyadai, Tsukuba, 305-0074 Japan
          phone: 81-298-36-9189, fax: 81-298-36-9199
          e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector      : pBACe3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI.
FEATURES  Location/Qualifiers
          source            1..1241
                           /organism="Mus musculus molossinus"
                           /mol_type="genomic DNA"
                           /sub_species="molossinus"
                           /db_xref="taxon:57486"
                           /clone="MSMg01-332A17.T7"
                           /sex="male"
                           /tissue_type="mixture of kidney and spleen"
                           /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match      2.4%; Score 48.4; DB 9; Length 1241;
Best Local Similarity 32.3%; Pred. No. 0.44;
Matches 155; Conservative 0; Mismatches 322; Indels 3; Gaps 1;
QY 715 TTTTAAAAAGAGCGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGCCA 774
Db 408 TTTTAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467
QY 775 GAAACGGTTGAAACCAACAGTACCACTGCGCAGGAACTAAGCGCGGAGAACTTCAAACT 834
Db 468 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 527
QY 835 AAAAAAGAGAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCA 894
Db 528 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 587
QY 895 CCAGAGGACTGGATGATGATGCGCAGACAGTGTACATTCAATGATGCTCAACAGGT 954
Db 588 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 647
QY 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTACTAGCCAGA---ACC 1011

```





```
Db      876  AAGGTGTGATWAGRAAAATTTTWWKTKRTTAARTGGTGATAARWAAAAAAGWAKGA 817
Qy      380  ACTGCCATGTACTATTTGGAGGAAAGGACTTTTAGTCAAGCTCAAGGGAATGTGTGAGAA 439
Db      816  RKGG---AGAAAAATGKGTGWRAAAAAAAGGCTGGAAGGCTGGAATTTTGTGT 761
Qy      440  GGCACATAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCCTGTATGTGCAACTAACAC 499
Db      760  GTAAAAAATATGTRTARWATAWATTTTAKKTKGTATGRTAARAWAAAAATATAAAT 701
Qy      500  CAGCTCAAGAGAAATTAACCTAAGAGAAATAGCAGACAGCAATGAGTGGTTACTCTACTTA 559
Db      700  AWAYGAATGAAAAAARAAAAAATGTRAWATAWATTTTAKKTKGTATGGAATWGTGK 641
Qy      560  CTTTATAGCATAGCAAAACCAAAAAAGACTTATACCAAGTGTGTCTTTTGGAAAAATGA 619
Db      640  GTKAWAAAAAATAAAAAAAGWAAWTAWKTARAAWDGKTKTKTTTTRAAWAAKGR 581
Qy      620  TTGCTTACTATTTTAACTAAAAAATAAGACTAGTCCACCAAGAGCGGAGGCT 679
Db      580  GTGKGAAGTTRDDWAAAAAATAAAAAAAGGCTGGAADARAAKGDAAARAD 521
Qy      680  ATTTCTTACGACTGCTG--GCTGGAAGAACTACTTTTAAAGAGCGGCGGCCA 737
Db      520  AGRWAAAAAKRTATATGTTGKTKTKGTAGTAAAGAKTAAATWAAAAAGTTGTAKWAT 461
Qy      738  TCTAGTGAGCAACTATACACTGATGACATGCGGCCAGAAACGGTTGAAACACACATAAC 797
Db      460  ATTAAAAAATAAAAAAAGKWTGRTGKGAATTTATRTAKKTTAAAAAATATGK 401
Qy      798  CACTGCGCAGGAACCTAAGCGCGCAGAAATTCAACTAAAAAAGAGTTCTTATTAAC 857
Db      400  WTRATKTKTKTGDKAARAAADAAATTAATAAAAAAATAAAAAAATAAAAAA 341
Qy      858  TACACTTAAAGAGCTGGTCATAAAGAGTAACCTCACCAGAGGCTGGATGATGCA 917
Db      340  AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 281
Qy      918  GCACAGACAGTTACATTTGAAATGATGCTCAACAGGCTGGAGAAACCTGCTGAAAAATAC 977
Db      280  AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 221
Qy      978  GTAGAGATTTGTACACTACTAGCCAGAACCAACAGCATTTGACTTATTTTGA 1037
Db      220  AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 161
Qy      1038  AAAAGCTGAACAGCAACCACTAACCACTTTTCACTGCTGCACACAGAGCCTGCAGAT 1097
Db      160  NNGNANNNANNANNAANNAANNAANNAANNAANNAANNAANNAATTHWCNNCWAATCTTT 101
Qy      1098  TTTTGTCTTTTCATGCTGGAACTATGTTAAAGTTTGCCATGCTATTTTGTGTTTTTAA 1157
Db      100  TATTASWTTNNNNNNNANANATWNTANNNNNNNNNNNNNNNNNNNNNNNNNN 41
Qy      1158  CAGACAAGGAGGCAAAAGAA 1178
Db      40  NNNNANNNNNCACCAAWTA 20
```

```
RESULT 14
CG753732/c
LOCUS
DEFINITION
P048-4-G03.ya Ppa EcorI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION
CG753732
VERSION
CG753732.1 GI:37978509
KEYWORDS
Pristionchus pacificus
SOURCE
Pristionchus pacificus
ORGANISM
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE
1 (bases 1 to 1811)
```

```
AUTHORS
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
22835951
12884007
COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Classes: BAC ends.

FEATURES
source
1..1811
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcorI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcorI and cloning into the BAC
vector."
```

Query Match 2.3%; Score 47.2; DB 9; Length 1811;  
Best Local Similarity 34.1%; Pred. No. 0.99;  
Matches 352; Conservative 0; Mismatches 677; Indels 2; Gaps 1;

Qy 36 AGCAACCACTCGTTTAAAGGAAAAAGTAAACAGGAAGTGTCTCTCATTTGTTTTAAAAA 95  
Db 1472 ANAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 1413

Qy 96 TGAATGTTCACTGAATGGAAAGATATCGGTGGAATAGTTACAAAAAGAGCTGCA 155  
Db 1412 AAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 1353

Qy 156 GGAGGACGAGCTGAAATCTTTACACGAGGAGCGGAAACCTACTTGGGACCAAGCGGGA 215  
Db 1352 NNANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 1293

Qy 216 CATGGAATGGGAAACCAACAGTGGATGAATGACCAAAAAAGCAAGTATTCATTTTGATTC 275  
Db 1292 AAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 1233

Qy 276 TTTGGTTAAAAATGTTTATTTTGAAGTCTTTACACAAAGAAATATATTTCTCTGCTGATGT 335  
Db 1232 NANAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 1173

Qy 336 TAATTGGTTTGTGCAACATGAATGGGAAAAAGACC--AAGGCTGGCACTGCCATGTACTA 393  
Db 1172 AAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 1113

Qy 394 ATTGGAGGAAAGACTTTTAGTCAAGCTCAAGGAAATGTGGAGAGGCAACTAAATGTT 453  
Db 1112 AAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 1053

Qy 454 TACTGGAGCAGATGGTTGGTAAACAGCCTGTATGTGCAACTAACACCAGCTGAAAGATT 513  
Db 1052 AAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 993

Qy 514 AAACCTAAGAGAAATAGCAGACGACATGAGTGGTTACTCTACTTACTTATAGCATAG 573  
Db 992 ANAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 933

Qy 574 CAAACCAAAAAAGACTTATACCAAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTT 633  
Db 932 NAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 873

Qy 634 TTAACCTAAAAAGAAATAGAGCACTAGTCCACAGAGAGCGGCTATTTTCTTACAGT 693



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:18:38 ; Search time 99 Seconds  
(without alignments)  
2435.011 Million cell updates/sec

Title: US-10-069-056-10

Perfect score: 3637

Sequence: 1 MAGNAYSDEVLGATNLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3627	99.7	672	4	AAY72706 Parvoviru
2	3623	99.6	672	4	AAY72702 Parvoviru
3	3620	99.5	672	4	AAY72704 Parvoviru
4	3618	99.5	672	4	AAY72708 Parvoviru
5	3618	99.5	672	4	AAY72710 Parvoviru
6	2371	65.2	690	1	AAP40306 Sequence
7	458.5	12.6	550	3	AAY58164 Adeno ass
8	458.5	12.6	550	3	AAY58169 Adeno-ass
9	458.5	12.6	610	3	AAY58159 Adeno ass
10	458.5	12.6	610	4	AAY97720 Rep prote
11	458.5	12.6	610	5	AAU11404 Adeno-ass
12	458.5	12.6	610	5	AAE22887 Adeno-ass
13	458.5	12.6	610	5	AAE22864 Adeno-ass
14	458.5	12.6	610	5	AAE26940 Adeno ass
15	458.5	12.6	610	6	ABU64865 Rep prote
16	458.5	12.6	610	6	ABU64760 Adeno ass
17	458.5	12.6	610	6	ABR43398 Adeno-ass
18	458.5	12.6	610	7	ADI40280 Adeno-ass
19	458.5	12.6	610	8	ADH58895 Predeterm
20	455.5	12.5	626	2	AAE5384 Barbary d
21	455.5	12.5	626	4	AAY97724 NSI prote
22	455.5	12.5	626	5	AAE22891 Muscovy d
23	455.5	12.5	626	5	AAE28648 Muscovy d
24	455.5	12.5	626	5	AAE26944 Muscovy d
25	455.5	12.5	626	6	ABU64869 NSI prote

26	455.5	12.5	626	6	ABU64764 Parvoviru
27	455.5	12.5	626	6	ABR43402 Muscovy d
28	455.5	12.5	626	7	ADI40288 Muscovy d
29	455	12.5	627	4	AAY97722 Rep prote
30	455	12.5	627	5	AAE22889 Barbare
31	455	12.5	627	5	AAE228646 Barbare
32	455	12.5	627	5	AAE26942 Barbare
33	455	12.5	627	6	ABU64867 Rep prote
34	455	12.5	627	6	ABU64762 Parvoviru
35	455	12.5	627	6	ABR43400 Barbare
36	455	12.5	627	7	ADI40284 Barbare
37	445.5	12.2	625	7	ABR62760 Adeno ass
38	437.5	12.0	536	2	AAW46311 AAV4 Rep
39	437.5	12.0	536	6	ABG73938 Adeno-ass
40	437.5	12.0	623	2	AAW46307 AAV4 Rep
41	437.5	12.0	623	2	AAW46312 AAV4 Rep
42	437.5	12.0	623	4	AAY97712 Rep prote
43	437.5	12.0	623	5	AAE28636 Adeno-ass
44	437.5	12.0	623	5	AAE26932 Adeno ass
45	437.5	12.0	623	6	ABU64857 Rep prote

#### ALIGNMENTS

#### RESULT 1

AAAY72706

ID AAY72706 standard; protein; 672 AA.

XX AC AAY72706;

XX AC AAY72706;

DT 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX XX

DE Parvovirus non-structure protein 1 (NS1) variant (T363A).

KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; variant.

XX OS

OS Parvovirus.

OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT Misc-difference 363

FT FT /note= "Wild type Thr substituted with Ala"

XX XX

PN EP1077260-A1.

XX PD 21-FEB-2001.

XX PF 13-AUG-1999; 99EP-00115161.

XX PR 13-AUG-1999; 99EP-00115161.

XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX PI Nuesch J, Rommelaere J;

XX XX WPI: 2001-212717/22.

XX DR N-PSDB; AAD02801.

XX PT Novel parvovirus non-structure protein variant, useful for treating

XX PT tumoral diseases, has a shifted equilibrium between DNA replication and

XX PT transcription activities, and cytotoxic activity.

XX PS Claim 6; Page 19-21; 41pp; English.

XX CC The present sequence is parvovirus non-structure protein 1 (NS1) variant  
(T363A). The invention relates to the variants of the parvovirus non-  
structure protein (NS1) having a shifted equilibrium between the DNA  
replication and transcription activities, and the cytotoxicity activity.  
These variants are useful as vectors for treating tumoural diseases. The  
variant DNAs are useful for gene therapy. (Updated on 06-AUG-







QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKEVSIKTLKELVHKVTSPE 300  
|||  
Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKEVSIKTLKELVHKVTSPE 300  
|||  
QY 301 DNMMPDSDSYENMAOPGGNLLKNLTLEICTLTARTKTAFLDILILEKAETSKLTNFSLPD 360  
|||  
Db 301 DNMMPDSDSYENMAOPGGNLLKNLTLEICTLTARTKTAFLDILILEKAETSKLTNFSLPD 360  
|||  
QY 361 TRACRIFAFHGMNVYKCHAIACCVLNRRQGGKRNVLFGHPASTGKSIIAQAIQAQAVGNVG 420  
|||  
Db 361 TRACRIFAFHGMNVYKCHAIACCVLNRRQGGKRNVLFGHPASTGKSIIAQAIQAQAVGNVG 420  
|||  
QY 421 CYNAAVNPPFNDCTKNLIWVEEAGNFGQVQVNOFAICSGQIRIDQKGGSKQIEPTP 480  
|||  
Db 421 CYNAAVNPPFNDCTKNLIWVEEAGNFGQVQVNOFAICSGQIRIDQKGGSKQIEPTP 480  
|||  
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540  
|||  
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540  
|||  
QY 541 NGYQSTMASYCAKWKVPDMSNWAEPKVPPTINLLGSARSPPTTPKSTPLSQNYALTPL 600  
|||  
Db 541 NGYQSTMASYCAKWKVPDMSNWAEPKVPPTINLLGSARSPPTTPKSTPLSQNYALTPL 600  
|||  
QY 601 ASDLEDLALEPWPSTPVPVAGTAETONTGEGAGSKACODGOLSPTWSEIEDLRACTGAEP 660  
|||  
Db 601 ASDLEDLALEPWPSTPVPVAGTAETONTGEGAGSKACODGOLSPTWSEIEDLRACTGAEP 660  
|||  
QY 661 LKDFSEPLNLD 672  
|||  
Db 661 LKDFSEPLNLD 672  
|||

## RESULT 6

AA040306  
ID AAP40306 standard; protein; 690 AA.

XX AC

XX AC

XX AC

XX AC

DT 24-OCT-2003 (revised)

DT 12-JAN-1992 (first entry)

XX Sequence encoded by the double-stranded replicative form DNA of porcine parvovirus.

XX Protein envelope; immunogen; vaccine; antigen; epitope.

XX Porcine parvovirus; NADL-2 virulent strain.

XX WO8402847-A.

XX 02-AUG-1984.

XX 19-JAN-1984; 84WO-US0000063.

XX 19-JAN-1983; 83US-00459203.

XX 06-JAN-1984; 84US-00567968.

XX (AMGE-) AMGEN.

XX FOX GM;

XX WPI; 1984-201354/32.

XX N-PSDB; AAN40252.

XX Polypeptide obd. by recombinant DNA methods - for vaccination against parvovirus infections in man and animals.

XX Claim 6; Table II, Page 33-49; 80pp; English.

XX The inventors claim an immunologically active polypeptide for the development of vaccinal immunity against parvovirus infection. Also

CC claimed are DNA sequences wholly or partly duplicative of defined sequences. The polypeptides are used in vaccines for conferring CC protection against parvovirus infections in man and animals. (Updated on CC 24-OCT-2003 to standardise OS field)

XX SQ Sequence 690 AA;

Query Match 65.2%; Score 2371; DB 1; Length 690;

Best Local Similarity 67.7%; Pred. No. 3.2e-220;

Matches 448; Conservative 70; Mismatches 128; Indels 16; Gaps 6;

QY 2 AGNAYSDVLTATNWLKEKSNQBFVSFKENNVOLNGKDIGNWSYKKELOEDELKSLQR 61  
|||  
Db 5 AGNTYSEELKATNWLQDNAQEAESYVFTQKVLNGLKEIAWNNTKDTTDEMINLQR 64  
|||  
QY 62 GAETTWQSDSEMEWETTVDEMTKKQVFI FDSL VKCLFEVLNTKNI FPGDVNMFVQHEWG 121  
|||  
Db 65 GAETSDQATDMEWESIDSLTKRQVLI FDSL VKCLFEGLKQLSPSDCYWFLQHEHG 124  
|||  
QY 122 KQGWCHVLIGGKDFSQAGKWRRLQNLVWRSRLVTACNVOLTFAERIKLREIADNE 181  
|||  
Db 125 QDTGYHCHVLLGGKGLQAGMWKRLQNLNLSRWLNQCKVPLTPVERIKJRELAEDGE 184  
|||  
QY 182 WTLTLYKKKOTKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGKWNFLK 241  
|||  
Db 185 WSLTYTHKQTKKQTKWTHFGNMIAYYFLNKKRKTTE--REHGYLSSDSGFWNFLK 242  
|||  
QY 242 EGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKEVSIKTLKELVHKVTSPE 301  
|||  
Db 243 EGERHLVSHLFTANKPETVEITVTTAQEVPRGRIQTKKEVSIKTI RDLVNRKCTSIEG 302  
|||  
QY 302 WMMQPDSDSYENMAOPGGNLLKNLTLEICTLTARTKTAFLDILILEKAETSKLTNFSLPD 361  
|||  
Db 303 WSMTPDSDSYENMAOPGGNLLKNLTLEICTLTARTKTAFLDILILEKAETSKLTNFSINT 362  
|||  
QY 362 RACRIEAFHGMNVYKCHAIACCVL--NRQGGKENTVLFHGPASTGKSIIAQAIQAQAVGNV 419  
|||  
Db 363 RTCKIFSMENNVYIKVCHAITCVLKQTRRKKYNSISCHGPASTGKSIIAQHIANLVGNV 422  
|||  
QY 420 GCYNAANVPPFNDCTKNLIWVEEAGNFGQVQVNOFAICSGQIRIDQKGGSKQIEPT 479  
|||  
Db 423 GCYNAANVPPFNDCTKNLIWVEEAGNFGQVQVNOFAICSGQIRIDQKGGSKQIEPT 482  
|||  
QY 480 PVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLV 539  
|||  
Db 483 PVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLV 542  
|||  
QY 540 KNGYQSTMASYCAKWKVPDMSNWAEPKVPPTINLLGSARSPPTTPKSTPLSQNYALTTP 599  
|||  
Db 543 KNGYQATMASYMHWHGNVDPWSEKLEPEKMHSPINTFTDSQIS--TSVKTSPADINYAATP 601  
|||  
QY 600 LASDLE--DLALEPWPSTPVPVAGTAETONTGEGAGSKACODGOL---SPTWSEIEEDLRAC 655  
|||  
Db 602 IQEDLDLALALEPWPSEPTTFTNHLHTTP-----PDSAIRTPSPWSEIETDIRAC 654  
|||  
QY 656 FG 657  
|||  
Db 655 FG 656  
|||

## RESULT 7

AA058164

ID AAY58164 standard; protein; 550 AA.

XX AC

XX AC

XX AC

DT 07-MAR-2000 (first entry)

XX Adeno associated virus AAV5 Rep68 protein.

XX Adeno associated virus; AAV5; AAV2: inverted terminal repeat; ITR; promoter; Rep protein; capsid protein; regulation; transcription; ITR; replication; chromosomal integration; tissue tropism; cellular receptor;





CC associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus  
CC which relies on a helper virus for replication, in the absence of which  
CC the AAV5 genome is integrated into a host chromosome in a locus specific  
CC manner. The method provides a way to deliver a nucleic acid to a specific  
CC regions, tissues and cell types of the central nervous system comprising  
CC inserting the nucleic acid between a pair of AAV inverted terminal  
CC repeats or delivering an AAV5 particle containing a vector comprising the  
CC nucleic acid. The method is useful for treating brain disorders such as  
CC demyelination disease, Alzheimer's disease and Parkinson's disease, and  
CC metabolic disorders such as musculoskeletal diseases, cardiovascular  
CC disease, cancer and autoimmune disorders, for treating genetic diseases  
CC such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoadosteronism,  
CC immotile cilia syndrome, and for treating bronchitis, pneumonia,  
CC emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is  
CC useful for delivering gene that may have a systematic effect like anti-  
CC hypertension drugs, insulin, coagulation factors, antibiotics, growth  
CC factors and hormones. This is the amino acid sequence of the adeno-  
CC associated virus 5 (AAV5) Rep68 protein, one of 4 Rep proteins that  
CC regulate replication and transcription of the AAV5 genome, described in  
CC the method of the invention  
XX  
SQ Sequence 550 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 550;  
Best Local Similarity 27.4%; Pred. No. 1.2e-34;  
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;  
67 WQSEDMEWETTVDENTKKQVFIFFDSLKKLFEVLNTRKNIFPDGVNMFVQHEWKGKQGW 126  
35 WELPPESDLNLTUVE--QPQLTVADIRRVFLYE---WNKFSQESKFFVQFEKGSY-F 88  
127 HCHVLIGGKDF-SQAQKGMWRRLNVNYSRWLTACNVQLTPAERIKLRBI-AEDNEWVT 184  
89 HLHVLVETSGISSMWLGRY-----VSQIRALQV---KVVFGQIEPQINDWA 132  
185 LLTYKHQKQKDYTKCVLFGNMIAYFLTKKLISTPPRGGYFLSDSGWKNFLKEGE 244  
133 I-----TKVKKGGANKVDSGYIPAYLLPK-----VOPELQAWNTLDE-- 171  
245 RHLVSKLYTDDMPETVETVTVAQETKGRIRITKKEVSIKTLKE-----LV 292  
172 -YKLAALNLEERKELVAQFLAESQSQEAAQO--REFSADPVIKTSQKYNALVNLV 228  
293 HKRVTSPEDMMQPDYSIEMMAQPGGENLLKNTLICTLTLARTKTAFLDILEKAETSK 352  
229 EHGITSQKQWIOBNSQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY----- 279  
353 LTNFSLPD-----TRACIFAFHGMN--YV-KVCHAICCVLNROGGKRNITVLFHGPASTGK 405  
280 LVGSSVPEDISKRNIRWQIFEMNGYDPAYAGSILYGWC---QRSFNKRNITVLYGPATTGK 336  
406 SIITAAIAQAVGNVGCYNAANVFPNDCTNKLIVVEAGNFGQOVNQPKAICSGOTIR 465  
337 TNIAEALHTVPPYGCUNVTNENFPNDQVDMKLIIWEEGKMTNKVSVESKALIGSKVR 396  
466 IDQKSGSKQIEPTVMTTNENITVYIGCEERPEHTQPIRDRMLNHLTHLPDGFGL 525  
397 VDQKCKSSVQIDSTPVIIVTNTNMCVVVDGNTSTFEHQPLEDRMFKFELTKLPLPDFGK 456  
526 VDKNEHFMICAWLVKNGYQSTMASYCAKWKGVDPDSE-----NWAEPKVPTPI-- 573  
457 ITRQEVKDFPAW-----AKVNOQVPVTHFEKVPRELAGTKGAESKLRPLGSD 502  
574 -----NLIGSARSPT--TPKSTPLSONYA-LTPL 600  
503 VTNVTSYKLEKRLARLFSVPTPRSSDVTVDPAERPL 539

RESULT 9  
AAV58159  
ID AAV58159 standard; protein; 610 AA.  
XX  
AC AAV58159;

XX 07-MAR-2000 (first entry)  
XX Adeno associated virus AAV5 Rep78 protein.  
XX  
XX Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR;  
XX promoter; Rep protein; capsid protein; regulation; transduction;  
XX replication; chromosomal integration; tissue tropism; cellular receptor;  
XX gene therapy; neutralising antibody; erythroid progenitor cell;  
XX transduction; cancer; genetic disease; Rep78.  
XX Adeno-associated virus 5.  
XX WO9961601-A2.  
XX 02-DEC-1999.  
XX 28-MAY-1999; 99WO-US011958.  
XX 28-MAY-1998; 98US-0087029P.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Chiorini JA, Kotin RM;  
XX WPI; 2000-062707/05.  
XX N-PSDB; AA249211.  
XX Adeno-associated virus 5 based vectors and particles, useful for gene  
XX therapy.  
XX Claim 15; Fig 6; 91pp; English.

This sequence represents the Rep78 protein of adeno associated virus type 5 (AAV5). The invention relates to vectors comprising a pair of AAV5 inverted terminal repeats (ITRs) with a promoter between the ITRs. The vector may comprise the viral genome, or subregions thereof, including sequences encoding Rep proteins and capsid proteins, and is encapsidated in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613), Rep52 (AAV58168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in regulation of replication and transcription, in addition to the production of progeny genomes. Rep68 and Rep78 are also associated with the stable integration of the viral genome into human chromosomes. The three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3 (AAV58162) assemble to form an icosahedral capsid, and differ from each other by the use of alternative splicing and an unusual translation initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2 capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are likely to utilise distinct cellular receptors and are serologically distinct. In a gene therapy application, therefore, AAV5 would allow for transduction of a patient who already possess neutralising antibodies either as a result of natural immunological defence or from prior exposure to AAV2 vectors. The vectors may be useful for transducing erythroid progenitor cells or cells lacking heparin sulphate proteoglycans, which is very inefficient with AAV2-based vectors. The vectors may also be useful for transducing cells with a nucleic acid of interest in order to produce cell lines that could be used to screen for agents that interact with the gene product of the nucleic acid of interest. In addition to transduction of other cell types, transduction of erythroid cells would be useful or the treatment of cancer and genetic diseases which can be corrected by bone marrow transplants using matched donors

SQ Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 3; Length 610;  
Best Local Similarity 27.4%; Pred. No. 1.5e-34;  
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQSEDMEWETTVDENTKKQVFIFFDSLKKLFEVLNTRKNIFPDGVNMFVQHEWKGKQGW 126  
Db 35 WELPPESDLNLTUVE--QPQLTVADIRRVFLYE---WNKFSQESKFFVQFEKGSY-F 88

QY 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLVACNVQLTPAERIKLREI-AEDNEWVT 184  
 Db 89 HLHVLVETSGISSMVLGRY-----VSQIRALV---KVVFQIEPQINDWA 132  
 QY 185 LLYYKHQTKDYTKCVLFGNMIAYFLTKKXISTSPRGGYFLSDSGWKNFLKEGE 244  
 Db 133 I-----TKVRKGGANKVVDGYIPAYLLPK-----VOPELQAWTNLDE-- 171  
 QY 245 RHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTLKE-----LV 292  
 Db 172 -YKLAALNLEERKRLVAQFLAESQSQEASQ--REFSADPVIKSKTSQKYMALVNLV 228  
 QY 293 HKRVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTLARTKTAFLILEKAETSK 352  
 Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY----- 279  
 QY 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNROGGKNTVLFHGPASTGK 405  
 Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGWC---QRSFNKNTVMYLGPAATGK 336  
 QY 406 SIITAOIAQAVGNVGCYNAANVFPNDCTKNLIWVEEAGNFQOQVNOQKALCSCGOTIR 465  
 Db 337 TNIAEAIAHTVPYGCNVNTNENFPNDCVDKMLIWEESGKMTNKVSESALIGGSKVR 396  
 QY 466 IDQKGSKQIEPTPVIMTNNENITVVRIGCEERPEHTQPIRDMNLNIHLTHLPGDFGL 525  
 Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHQQPLEDRMFKFELTKLPDPFGK 456  
 QY 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NNAEPKVPPI-- 573  
 Db 457 ITKQEVKDFPFW-----AKVQVPTVTHEFKVPRELAGTKGAEKSLKRLPLGD 502  
 QY 574 -----NLGARSPT--TPKSTPLSQNYA-LTPL 600  
 Db 503 VTNTSYKSLKRALSFVPETPRSSDVTVDPAFLRPL 539

## RESULT 10

AAV97720 standard; protein; 610 AA.

XX AC AAV97720;

DT 19-JUN-2001 (first entry)

XX DE Rep protein sequence.

XX KW Fusion nucleic acid library; Rep protein; tumour cell; apoptosis;  
 KW nucleic acid modification enzyme; cell death; decreased cell growth;  
 KW protein-protein interaction detection; cell division; cancer therapy;  
 KW protein drug discovery; pharmacogenetics.

XX OS Adeno associated virus 5.

XX PN WO200114539-A2.

XX PD 01-MAR-2001.

XX PF 18-AUG-2000; 2000WO-US022906.

XX PR 20-AUG-1999; 99US-0150004P.

XX PR 02-JUN-2000; 2000US-0209130P.

XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

XX PI Li M;

XX DR WPI; 2001-218443/22.

XX DR N-PSDB; AAA91310.

XX PT New library of fusion nucleic acids each encoding a Rep protein  
 PT recognized by a nucleic acid modification enzyme and a candidate protein,  
 PT useful for detecting protein-protein interactions, protein drug discovery

or pharmacogenetics.

Disclosure; Fig 21; 106pp; English.

XX This sequence is the adeno associated virus 5 Rep protein. The invention  
 CC relates to a library of fusion nucleic acids, each encoding a Rep  
 CC protein, a candidate protein, a presentation structure, a targeting  
 CC sequence or a label. The Rep protein is a nucleic acid modification  
 CC enzyme. The random or directed libraries (including the cDNA libraries)  
 CC can be introduced into any tumour cell, and peptides identified which by  
 CC themselves induce apoptosis, cell death, loss of cell division or  
 CC decreased cell growth. The methods and compositions may also be used to  
 CC detect protein-protein interactions, protein drug discovery, particularly  
 CC for protein drugs that interact with targets on cell surfaces, to  
 CC discover DNA or nucleic acid binding proteins, using nucleic acids as  
 CC targets, to screen for nucleic acid modification enzymes with decreased  
 CC toxicity for the host cells, to identify or generate Rep proteins with  
 CC decreased toxicity, improved enzyme attachment sequences for use in  
 CC expression vectors and in pharmacogenetic studies. The method is useful  
 CC in cancer therapy and in killing tumour cells. The methods can be  
 CC combined with other cancer therapeutics (drugs or radiation) to sensitize  
 CC cells and thus induce rapid and specific apoptosis, cell death, loss of  
 CC cell division or decreased cell growth after exposure to a secondary  
 CC agent

XX SQ Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 4; Length 610;

Best Local Similarity 27.4%; Pred. No. 1.5e-34;  
 Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQOSEMEWETTVDENETKQVFIQFISLVKKLPEVLNKNIFPGDVNMFVQHEWGDQGW 126  
 Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSQESKFVFOFEKGEY-F 88  
 QY 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLVACNVQLTPAERIKLREI-AEDNEWVT 184  
 Db 89 HLHVLVETSGISSMVLGRY-----VSQIRALV---KVVFQIEPQINDWA 132  
 QY 185 LLYYKHQTKDYTKCVLFGNMIAYFLTKKXISTSPRGGYFLSDSGWKNFLKEGE 244  
 Db 133 I-----TKVRKGGANKVVDGYIPAYLLPK-----VOPELQAWTNLDE-- 171  
 QY 245 RHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTLKE-----LV 292  
 Db 172 -YKLAALNLEERKRLVAQFLAESQSQEASQ--REFSADPVIKSKTSQKYMALVNLV 228  
 QY 293 HKRVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTLARTKTAFLILEKAETSK 352  
 Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY----- 279  
 QY 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNROGGKNTVLFHGPASTGK 405  
 Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGWC---QRSFNKNTVMYLGPAATGK 336  
 QY 406 SIITAOIAQAVGNVGCYNAANVFPNDCTKNLIWVEEAGNFQOQVNOQKALCSCGOTIR 465  
 Db 337 TNIAEAIAHTVPYGCNVNTNENFPNDCVDKMLIWEESGKMTNKVSESALIGGSKVR 396  
 QY 466 IDQKGSKQIEPTPVIMTNNENITVVRIGCEERPEHTQPIRDMNLNIHLTHLPGDFGL 525  
 Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHQQPLEDRMFKFELTKLPDPFGK 456  
 QY 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NNAEPKVPPI-- 573  
 Db 457 ITKQEVKDFPFW-----AKVQVPTVTHEFKVPRELAGTKGAEKSLKRLPLGD 502

RESULT 11

AAU11404	127	HCHVLIGGKDP-SQAQGWRRQLNYYWRSWLTACNVOLTPAERIKLREI-AEDNEWVT	184
ID	AAU11404	standard; protein; 610 AA.	
AC	AAU11404;		
XX			
DT	26-FEB-2002	(first entry)	
XX			
DE	Adeno-associated virus 5 (AAV5), Rep78 protein.		
XX			
KW	Adeno-associated virus 5; AAV5; Rep78; neurotropic; neuroprotective;		
KW	cytostatic; gene therapy; Parkinson's disease; Alzheimer's disease;		
KW	demyelination disease; metabolic disorder; musculoskeletal disease;		
KW	cardiovascular disease; cancer; autoimmune disorder; genetic disease;		
KW	cystic fibrosis; pseudohypoadosteronism; immotile cilia syndrome;		
KW	bronchitis; pneumonia; emphysema; pulmonary oedema;		
KW	central nervous system; replication; transcription.		
XX			
OS	Adeno-associated virus 5.		
XX			
PN	WO200170276-A2.		
XX			
PD	27-SEP-2001.		
XX			
PF	22-MAR-2001; 2001WO-US009123.		
XX			
PR	22-MAR-2000; 2000US-00533427.		
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
XX			
PI	Chiorini JA, Kotin RM, Davidson B, Zabner J;		
XX			
DR	WPI; 2002-055104/07.		
DR	N-PSDB; AAS17712.		
XX			
PT	Delivering nucleic acid into cell for treating Parkinson's disease, by		
PT	administering to cell an adeno-associated virus 5 particle comprising the		
PT	nucleic acid inserted between a pair of AAV inverted terminal repeats.		
XX			
PS	Disclosure; Fig 6; 130pp; English.		
CC	The invention describes a novel method of delivering a nucleic acid into		
CC	a cell in a subject, comprising administering to the cell an adeno-		
CC	associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus		
CC	which relies on a helper virus for replication, in the absence of which		
CC	the AAV5 genome is integrated into a host chromosome in a locus specific		
CC	manner. The method provides a way to deliver a nucleic acid to a specific		
CC	regions, tissues and cell types of the central nervous system comprising		
CC	inserting the nucleic acid between a pair of AAV inverted terminal		
CC	repeats or delivering an AAV5 particle containing a vector comprising the		
CC	nucleic acid. The method is useful for treating brain disorders such as		
CC	demyelination disease, Alzheimer's disease and Parkinson's disease, and		
CC	metabolic disorders such as musculoskeletal diseases, cardiovascular		
CC	disease, cancer and autoimmune disorders, for treating genetic diseases		
CC	such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoadosteronism,		
CC	immotile cilia syndrome, and for treating bronchitis, pneumonia,		
CC	emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is		
CC	useful for delivering gene that may have a systematic effect like anti-		
CC	hypertension drugs, insulin, coagulation factors, antibiotics, growth		
CC	factors and hormones. This is the amino acid sequence of the adeno-		
CC	associated virus 5 (AAV5) Rep78 protein, one of 4 Rep proteins that		
CC	regulate replication and transcription of the AAV5 genome, described in		
CC	the method of the invention		
XX			
SQ	Sequence 610 AA;		
Query Match	12.6%; Score 458.5; DB 5; Length 610;		
Best Local Similarity	27.4%; Pred. No. 1.5e-34;		
Matches	158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;		
QY	67	WDQSEDEWETTVDEMTKQVFFDLSLVKKCLPELVNTKIFDGVNWFVQHWGKDGQW	126
DB	35	WELPPESDLNLTVE--QPQLTVADIRRRVFLYE---WNKFSQESKFFVQFEKGSY-F	88

Library of nucleic acid/protein conjugates, has a fusion of nucleic acid modification enzyme and candidate compound, and expression vector having

a fusion of nucleic acids encoding NAM enzyme and the compound.

Disclosure; Fig 21; 96pp; English.

The present invention relates to genetic libraries of nucleic acid/protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic acid modification (NAM) enzyme (B) and candidate compound), an expression vector (with a fusion of nucleic acids encoding the enzyme and candidate protein respectively), an enzyme attachment sequence (EAS; RNA sequence), where the candidate compound and candidate protein are different and EAS and the enzyme are covalently linked. The NAP conjugates are useful in screens to assay binding to target molecules and/or to screen candidate agents for the ability to modulate the activity of the target molecule. They are useful in cancer therapy. Sequences of the invention are also useful to detect protein-protein interaction, in drug discovery, to discover DNA or nucleic acid binding proteins, using nucleic acids as the targets and to screen for NAM enzymes with decreased toxicity for host cells (specifically Rep proteins with reduced toxicity). NAP conjugates are also useful in pharmacogenomic studies, for screening bioactive agents on surface cells, viruses and microbial organisms. They are also useful for screening proteins causing phenotypic changes such as overproduction or inhibition of protein expression, or proteins that alter attachment, infectivity, etc. of the virus. Sequences of the invention are also used in gene therapy. The present sequence is adeno-associated virus (AAV) 5 Rep. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 610;

Best Local Similarity 27.4%; Pred. No. 1.5e-34;

Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

```
PT CC The present invention relates to genetic libraries of nucleic acid/
XX CC protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic
PS CC acid modification (NAM) enzyme (B) and candidate compound), an expression
XX CC vector (with a fusion of nucleic acids encoding the enzyme and candidate
CC CC protein respectively), an enzyme attachment sequence (EAS; RNA sequence),
CC CC where the candidate compound and candidate protein are different and EAS
CC CC and the enzyme are covalently linked. The NAP conjugates are useful in
CC CC screens to assay binding to target molecules and/or to screen candidate
CC CC agents for the ability to modulate the activity of the target molecule.
CC CC They are useful in cancer therapy. Sequences of the invention are also
CC CC useful to detect protein-protein interaction, in drug discovery, to
CC CC discover DNA or nucleic acid binding proteins, using nucleic acids as the
CC CC targets and to screen for NAM enzymes with decreased toxicity for host
CC CC cells (specifically Rep proteins with reduced toxicity). NAP conjugates
CC CC are also useful in pharmacogenomic studies, for screening bioactive
CC CC agents on surface cells, viruses and microbial organisms. They are also
CC CC useful for screening proteins causing phenotypic changes such as
CC CC overproduction or inhibition of protein expression, or proteins that
CC CC alter attachment, infectivity, etc. of the virus. Sequences of the
CC CC invention are also used in gene therapy. The present sequence is adeno-
CC CC associated virus (AAV) 5 Rep. (Updated on 29-AUG-2003 to standardise OS
XX CC field)
SQ CC Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 610;
Best Local Similarity 27.4%; Pred. No. 1.5e-34;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQSEDMWETTVDEMTKKQVFIFDSLVKKCLFEVLNTKNIIFPGDVNMFVQHEMGKQGW 126
DB 35 WELPPSDMLNLTLVE--QPOLTVADRIRRVFLVE---WNKFSKQESKFFVQFKGSEY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVQLTTPAERIKLREI-AEDNEWVT 184
DB 89 HLHVLTVETSGISSVNLGRY-----VSQIRALV---KVVFQGLEQINDWA 132

QY 185 LLYTKHKQTKDYTKCVLFQGNMAYPLTKKISTSPRGGYFLSDSGWKNFLKEGE 244
DB 133 I-----TKVKKGGANKVVDGYPVALLPK-----VOPELQAWTNLDE-- 171

QY 245 RHLVSKLYTDDMPETVETTVTAQETKRGRIOTKKEVSIKTLKE-----LV 292
DB 172 -YKLAALNLEERKLVAQFLAESQSQEAAQO--REFSADPVIKSTSKYKVALVNLV 228

QY 293 HKEVTPEDMMQPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKATSK 352
DB 229 EHGITSEKQWIOENQESYLSFNSTGNSRSQIKAALDNATKIMSITSKSAVDY----- 279

QY 353 LTNFSLPD-----TRACRIPAFHGN--YV-KVCHATCCVLRNROGGKRNVTFLFGPASTGK 405
DB 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGWC---QRSFNKENTVWLYGPATTKG 336

QY 406 SIIAQATAQAVNGVGVNAANVFPNDCKNKLIVVEBAGNFGQQVQFKAICSGQITR 465
DB 337 TNIAEALAHVTPVGCNVNWNENFPNDCVDRKMLIWNWEGKMTKNVVEKAILGSKVR 396

QY 466 IDQKGKSGKOIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
DB 397 VDQKSSSVQIDSTPVIIVTNTNMCVVVDGNTSTTFEHHQPLEDRMPKFKLTLPDPFGK 456

QY 526 VDKNEWPMICAWLVNGYQSTMASYCAKWKGVDPWSE-----NWAEPKVPTPI-- 573
DB 457 ITRQEVKDFFAW-----AKVNOQVPTVTFEYKVPRELAGTKGAESLKRPLGD 502

QY 574 -----NLGARSRPET--TPKSTPLSONYA-LTPL 600
DB 503 VMTSYKSLKRALRSLFVPTPRSSDVTVDPAPLRPL 539
```

RESULT 13

AAE28644

ID AAE28644 standard; protein; 610 AA.

XX AAE28644;

AC AAE28644;

XX 29-AUG-2003 (revised)

DT 27-DEC-2002 (first entry)

XX Adeno-associated virus 5 Rep protein.

XX Nucleic acid modification enzyme; NAM; enzyme attachment sequence; EAS;

XX protein design automation; PDA; cancer; protein-protein interaction;

XX infection; gene therapy; Rep protein.

XX Adeno associated virus; 5.

OS WO200268453-A2.

XX 06-SEP-2002.

XX 19-FEB-2002; 2002WO-US004853.

XX 22-FEB-2001; 2001US-00792629.

XX (XENC-) XENCOR INC.

XX Li M, Dahiyat BI;

XX WPI; 2002-691653/74.

XX N-PSDB; AAD46138.

XX Generating a library of fusion nucleic acids for treating cancer or

XX infection, or detecting protein-protein interaction, comprises providing

XX computationally-derived library of candidate protein sequences and

XX expression vectors.

XX Disclosure; Page 180-182; 246pp; English.

XX The present invention relates to a novel method of generating a library

XX of fusion nucleic acids. The method involves providing a computationally-

XX derived library of candidate protein sequences and creating a library of

XX expression vectors containing a fusion nucleic acid having a sequence

XX encoding a nucleic acid modification (NAM) enzyme and a sequence encoding

XX a candidate protein sequence from the library and an enzyme attachment

XX sequence (EAS) that is recognised by the NAM enzyme. The invention also

XX relates to the use of a variety of computation methods including protein

XX design automation (PDA). The method is useful in generating and screening

XX fusion nucleic acids that may be used in treating cancer or infections,

XX in detecting protein-protein interactions, discovery of DNA or nucleic

XX acid binding proteins, protein drug discovery, screening for NAM enzymes

XX with decreased toxicity to the host cells and NAM enzyme/EAS pairs with

XX increased affinity or in pharmacogenetic studies. The invention is also

XX used in gene therapy. The present sequence is Adeno-associated virus 5

XX Rep protein. This sequence is used to illustrate the method of the

XX invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 610;

Best Local Similarity 27.4%; Pred. No. 1.5e-34;

Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQSEDMWETTVDEMTKKQVFIFDSLVKKCLFEVLNTKNIIFPGDVNMFVQHEMGKQGW 126

DB 35 WELPPSDMLNLTLVE--QPOLTVADRIRRVFLVE---WNKFSKQESKFFVQFKGSEY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVQLTTPAERIKLREI-AEDNEWVT 184

DB 89 HLHVLTVETSGISSVNLGRY-----VSQIRALV---KVVFQGLEQINDWA 132

QY 185 LLYTKHKQTKDYTKCVLFQGNMAYPLTKKISTSPRGGYFLSDSGWKNFLKEGE 244

Db 133 i-----TKVKKGGANKVDSGYIPAYLLPK-----VQPELQAWNTLDE-- 171  
Qy 245 RHLVSKLYTDDMRPEVETVTTTAQETKRGRIOTKKEVSIKTLKE-----LV 292  
Db 172 -YKLAALNLEERKRLVAQFLAESSQSQEASQ--REFSADPVIKTSQKYMALVNLV 228  
Qy 293 HKRVTSPEDEMMQPDYSIEMMAQPGENLLKNTLICTLTARTKTAFDLILEKAETSK 352  
Db 229 EHGITSKQWIOBQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY----- 279  
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNROGGKNTVLFHGPASTGK 405  
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC--QRSFNKNTVWLYGPATTGK 336  
Qy 406 SIQAQIAQAVNGVGCYNAANVFPNDCTNKNLIWVEAGNFGQOVNOFKATCSQTIR 465  
Db 337 TNIAEAIAHTVPYGCNVNTNENFPNDQVCKMLIWWEEGKMTNKVVESAKILGSKVR 396  
Qy 466 IDQKGGSKQIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525  
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNTTTFEHQOPLDRMFKFELTKRLPDPFGK 456  
Qy 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDSE-----NWAEPKVPTPI-- 573  
Db 457 ITKQEVKDFPFW-----AKNQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502  
Qy 574 -----NLGARSPT--TPKSTPLSONYA-LTPL 600  
Db 503 VTNYSKLEKRLARLSFVPETPRSSDVTVDPAERPL 539  
RESULT 14  
AAE26940  
ID AAE26940 standard; protein; 610 AA.  
AC AAE26940;  
XX  
DT 13-DEC-2002 (first entry)  
XX  
DE Adeno associated virus 5 Rep protein.  
XX  
KW Prokaryotic library; candidate protein; nucleic acid modification; NAM;  
KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;  
KW enzymology; cosmetic research; toxic; environmental safety assessment;  
KW nutrient biology; Rep protein.  
XX  
OS Adeno associated virus.  
XX  
PN WO200266653-A2.  
XX  
PN 29-AUG-2002.  
XX  
PF 14-DEC-2001; 2001WO-US049058.  
XX  
PR 14-DEC-2000; 2000US-0256163P.  
XX  
XX (XENC-) XENCOR INC.  
XX  
XX Li M, Liu Y;  
XX  
XX WPI; 2002-667068/71.  
DR N-PSDB; AAD44600.  
XX  
XX New library of prokaryotic pET-24a expression vectors, host cells or  
PT nucleic acid/protein conjugates, useful for screening candidate proteins  
PT and their nucleic acids or modification enzymes for pharmacogenetic  
PT analysis.  
XX  
XX Disclosure; Fig 21; 127pp; English.  
PS  
XX The invention relates to methods and compositions for the construction of  
CC prokaryotic libraries expressing candidate proteins and the use of these

CC libraries to identify candidate proteins and the nucleic acids encoding  
CC them. The invention provides a library of prokaryotic pET-24a vectors  
CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a  
CC nucleic acid modification (NAM) enzyme or a candidate protein, or a  
CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or  
CC the candidate protein, and an enzyme attachment sequence (EAS) recognised  
CC by the NAM enzyme. The library is used for identifying candidate proteins  
CC and nucleic acids encoding these proteins, in screening for NAM enzymes  
CC with decreased toxicity for the host cells, or in identifying novel or  
CC improved EASs, which may be used for understanding cellular processes or  
CC any subsequent therapeutic or toxic activities. The nucleic acid/protein  
CC (NAP) conjugates are useful in diagnostic assays and in research  
CC including clinical pharmacology, functional genomics, pharmacogenomics,  
CC agricultural chemicals, environmental safety assessment, chemical sensor,  
CC nutrient biology, cosmetic research or enzymology. These may also be used  
CC in vitro screening techniques and in assays with target molecules. The  
CC present sequence is Adeno associated virus 5 Rep protein used in the  
CC invention  
XX  
SQ Sequence 610 AA;  
Query Match 12.6%; Score 458.5; DB 5; Length 610;  
Best Local Similarity 27.4%; Pred. No. 1.5e-34;  
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;  
Qy 67 WQSEDEMEVETTVDEMTKQVFIQFDSLVKKCLPEVLNTKNIFPGDVNWFVQHEWGDQGW 126  
Db 35 WELPPESDLNLTIVE--QPQLTVADRIRRVFLYE---WNKFSQESKSFVQFQKSEY-F 88  
Qy 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVOLTPAERIKLRI-AEDNEWVT 184  
Db 89 HLHVLTVETSGISMWLGRY-----VSQIRAOQLV---KVVFQIEPQINDMVA 132  
Qy 185 LLYKHKQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSDSGSKNTFLKEGE 244  
Db 133 i-----TKVKKGGANKVDSGYIPAYLLPK-----VQPELQAWNTLDE-- 171  
Qy 245 RHLVSKLYTDDMRPEVETVTTTAQETKRGRIOTKKEVSIKTLKE-----LV 292  
Db 172 -YKLAALNLEERKRLVAQFLAESSQSQEASQ--REFSADPVIKTSQKYMALVNLV 228  
Qy 293 HKRVTSPEDEMMQPDYSIEMMAQPGENLLKNTLICTLTARTKTAFDLILEKAETSK 352  
Db 229 EHGITSKQWIOBQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY----- 279  
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNROGGKNTVLFHGPASTGK 405  
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC--QRSFNKNTVWLYGPATTGK 336  
Qy 406 SIQAQIAQAVNGVGCYNAANVFPNDCTNKNLIWVEAGNFGQOVNOFKATCSQTIR 465  
Db 337 TNIAEAIAHTVPYGCNVNTNENFPNDQVCKMLIWWEEGKMTNKVVESAKILGSKVR 396  
Qy 466 IDQKGGSKQIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525  
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNTTTFEHQOPLDRMFKFELTKRLPDPFGK 456  
Qy 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDSE-----NWAEPKVPTPI-- 573  
Db 457 ITKQEVKDFPFW-----AKNQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502  
Qy 574 -----NLGARSPT--TPKSTPLSONYA-LTPL 600  
Db 503 VTNYSKLEKRLARLSFVPETPRSSDVTVDPAERPL 539  
RESULT 15  
ABU64865  
ID ABU64865 standard; protein; 610 AA.  
XX  
AC ABU64865;  
XX  
DT 14-MAY-2003 (first entry)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:50 ; Search time 25 Seconds  
(without alignments)  
1782.628 Million cell updates/sec

Title: US-10-069-056-10  
Perfect score: 2637  
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKQDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	437.5	12.0	536	4	US-09-532-594B-10
2	437.5	12.0	623	4	US-09-532-594B-2
3	437.5	12.0	623	4	US-09-532-594B-11
4	431.5	11.9	546	4	US-09-807-802A-7
5	431.5	11.9	623	4	US-09-807-802A-2
6	431.5	11.9	623	4	US-09-807-802A-5
7	430.5	11.8	312	4	US-09-532-594B-8
8	430.5	11.8	399	4	US-09-532-594B-9
9	423.5	11.6	322	4	US-09-807-802A-11
10	423.5	11.6	399	4	US-09-807-802A-9
11	144	4.0	276	3	US-09-171-461-3
12	144	4.0	276	4	US-09-970-711-3
13	121.5	3.3	647	3	US-09-300-909-17
14	112	3.1	589	4	US-09-107-532A-6086
15	111	3.1	564	3	US-09-134-001C-2897
16	105	2.9	481	4	US-09-252-991A-30905
17	102	2.8	1170	4	US-09-638-524A-2
18	101	2.8	715	4	US-09-248-796A-20602
19	101	2.8	1120	4	US-09-792-024-95
20	100.5	2.8	2409	6	5180808-2
21	100	2.7	605	4	US-09-107-532A-6077
22	100	2.7	1503	4	US-09-677-046A-6
23	100	2.7	1509	4	US-09-677-046A-2
24	99	2.7	919	4	US-09-248-796A-19408
25	98.5	2.7	657	3	US-09-300-909-16
26	98	2.7	1343	4	US-09-270-767-31792
27	98	2.7	1343	4	US-09-270-767-47009

28	97.5	2.7	556	4	US-09-248-796A-18763	Sequence 18763, A
29	97.5	2.7	3248	1	PCT-US95-700-1	Sequence 1, Appl1
30	97.5	2.7	3248	5	PCT-US95-16216-1	Sequence 1, Appl1
31	96.5	2.7	957	4	US-09-489-039A-11233	Sequence 11233, A
32	96	2.6	291	4	US-09-107-532A-6390	Sequence 6390, Ap
33	96	2.6	422	3	US-09-134-001C-3034	Sequence 3034, Ap
34	95.5	2.6	441	4	US-09-540-236-2983	Sequence 2983, Ap
35	95	2.6	590	4	US-09-134-000C-4755	Sequence 4755, Ap
36	95	2.6	1307	1	US-08-395-246C-2	Sequence 2, Appl1
37	94	2.6	370	4	US-09-583-110-3405	Sequence 3405, Ap
38	94	2.6	1477	4	US-09-206-942-71	Sequence 71, Appl
39	94	2.6	1767	4	US-09-583-110-3975	Sequence 3975, Ap
40	94	2.6	2138	4	US-09-583-110-5274	Sequence 5274, Ap
41	94	2.6	15281	2	US-08-471-119A-2	Sequence 2, Appl1
42	93.5	2.6	506	4	US-09-540-236-2242	Sequence 2242, Ap
43	93.5	2.6	692	4	US-09-248-796A-19136	Sequence 19136, A
44	93.5	2.6	821	3	US-09-422-869-24	Sequence 24, Appl
45	93.5	2.6	821	4	US-09-538-092-972	Sequence 972, App

ALIGNMENTS

RESULT 1

US-09-532-594B-10  
; Sequence 10, Application US/09532594B  
; Patent No. 6468524  
; GENERAL INFORMATION:  
; APPLICANT: Chordini, John A.  
; APPLICANT: Kotin, Robert M.  
; APPLICANT: Safar, Brian  
; APPLICANT: Davidson, Beverly  
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF  
; FILE REFERENCE: 14014.0252U2  
; CURRENT APPLICATION NUMBER: US/09/532,594B  
; CURRENT FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 536  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =  
; OTHER INFORMATION: synthetic construct  
; NAME/KEY: misc feature  
; OTHER INFORMATION: AAV4 Rep protein 68  
US-09-532-594B-10

Query Match	12.0%	Score 437.5;	DB 4;	Length 536;
Best Local Similarity	27.7%	Pred. No. 1.7e-36;		
Matches 157;	Conservative 75;	Mismatches 232;	Indels 103;	Gaps 19;
QY	63	AETWQSDMEWETTVDENTKQVIFDLSLVKKCLFEVLTKNIFPGDVNFWVQHEWGK	122	
DB	31	AEKWELEPPDSMDLNLIE--QAPLTVAEKQLQREFLVEMRRVSK--APEALFFVQPEKG-	85	
QY	123	DOGMHCHVLIG-----GKDFSOAQGQKWRQLNVYMSRWLTACNVQLTPAERIKL	173	
DB	86	DSYFHLHLVETGVKSMVGVYVSQIKEK-----LVTRIYRGVEP-----	126	
QY	174	RETAEDNEWTLTYKHQ--TKQDYTKCVLFGNMIAIYFLTKKKIKISTSPRGGGYFLSS	231	
DB	127	-----QLPNPNFAVTKTRNGAGGNKVDDC-----YIPNYLLPK-----TQP-----	163	
QY	232	DSGMK--TNF-----LKEGERHLVSKLYTDDMPETVETVTTTAQETKGR-----	275	
DB	164	ELQWANTNMDOYISACINLAERKRLVAQHL-----THVSOTQEQNKQNPNSD	212	
QY	276	---IQTKKEVSIKTKLKLVLHKVTSPEWMMQPDYSIEMMAQPGGNNLLKKTLEICTL	332	
DB	213	APVIRKTSARYNELVGLVDRGKITSEKQWIQEDQASISFNASNSRSRQIKKALDNASK	272	

333	Qy	TLARTTAPDLILEKAEFTSLNFSLPDTRACRIAPFHGWNVYKVCHAI	CVVLNRGGKKR	392
273	Db	IMSLTTTAPDYLGVQNPDPEDIST-----NRYIRILEMNGYDPQYAA	SVFLGWAKQFFGKR	327
393	Qy	NTVLFHGHPASTGKSIITAAQIAQAVGNVCYNAANVPFNDCTNKNLI	WVEEAGNFGQV	452
328	Db	NTIWLFGPATTKTNIAEALAHAVPFYCGVNTWNTENPFND	CVDYDXVNIWVEEGKMTAKV	387
453	Qy	NQFKATCSGQTTIRIOKKGSKQIETPTVITNTNITVTVRIGCEBER	PEHTQPIRDRMLN	512
388	Db	ESAKAILGSKYRVQDKKSSAQIDPTFVIVTSTNTNCAVIDGNS	TTFEHHQQLQDRMFK	447
513	Qy	IHLTHLLPGDFGLVDKNEWPMICAWLKNVQGSMTASVCAKWG--	KVPDKSE--NWAEPKV	569
448	Db	FELTKLELHDGKVKYQEVKQDFRWASDHVTETVTFYVRKGARKR	PAPNDADISEPKR	507
570	Qy	PTPINLLGSARSFPTTPKSTPLSQNYA	596	
508	Db	ACP-----SVAQFSTSDAEAPV--	DYA	527

```

RESULT 2
US-09-532-594B-2
; Sequence 2, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc.feature
; OTHER INFORMATION: AAV4 Rep protein (full length)
US-09-532-594B-2

```

```

Query Match      12.0%; Score 437.5; DB 4; Length 623;
Best Local Similarity 27.7%; Pred. No. 2.2e-36;
Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;

QY      63 AETTDQSEDMEWETVDWMTKKQVFTFDSLVKKCLFEVLNTHKNTPPGDVNNWFWQHEWGK 122
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      31 AEKEWELPPDSMDMLNLTIE--QAPLTVAEKLQREFLVEWRRVSK--APEALFFVQFEKG- 85
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      123 DOGWCHVLIG-----CKDPSQAQGWRRQLNVYWSRWLYTACNVQLTPAERIKL 173
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      86 DSYFHLHLITVGVKSMWGVYVSQTKK-----LVTRIYRGVEP----- 126
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      174 REIADNEWVTLTAVKKHQ--TKDKYTKCVLFGNMIAIYFLTKKKISTSPRDGGYFLSS 231
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      127 ----OLPNWFAVTKRNGAGGKNKVDDC-----YIPNYLLPK---TOP----- 163
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      232 DSGWK-TNF-----LKEGERHLVSKLYTDMRPETVETVTVTTAQETKGR----- 275
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      164 ELQAWTWNMQYISACLNLAEKELVAQHL-----THVSQTQEQNKENQNPSND 212
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      276 ---IQTKKBSIKTTLKELVHKRVTSPEDDMMWQPSYIEMMAQPGENLLKNTLEICTL 332
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      213 APVIRSKTSARYMELVGWLVDRGITSEKQYIQEDQASYISFNAAANSRSQIKAAALDNASK 272
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      333 TLARTKTAFDLILKEAETSKLTFNPSLDDTRACRIFAFHGWNVVYKVCIAICCVLNRQGGR 392
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	273	IMSLTKTAPDYLVGNQPPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKFGKR	327
Qy	393	NTVLPHGPASTGKSIIAQIAAQVAGNVGCYNAANVFPFNDCTNKNLIWVEAGNFGQOV	452
Db	328	NTIWLFGPATTKTNIAEALIAHVPFYCCVNWNTNENFPFNDVCVKQVIWVEGKMTAKVV	387
Qy	453	NQFKAI CSGQTI RIDDQKSGKQIEPTPIVIMTNNENITVVRIGCSERPHEHTQPIRDRMLN	512
Db	388	ESAKAILGSGSVRVDDQCKSSAQIDPTPIVITSTNMCAVIDGNSNTEFHEHQOPLQDRMEK	447
Qy	513	IHLTHHLPFGDFGLVDKNEWPMICAWLVKNYGQSTWASCAKWG--KVPDWSS--NNAEPKV	569
Db	448	FELTKRLEHDFGKVTQBEVKDFPRWASHDVTEVTFEYVRKGGARKRPAPNDADISEPKR	507
Qy	570	PTPNILGARSPTTPKSTPLSQNYA	596
Db	508	ACP-----SVAQPSTSDAAEPV--DYA	527

```

RESULT 3
US-09-532-594B-11
; Sequence 11, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 78
US-09-532-594B-11

```

Query Match	12.0%	Score 437.5	DB 4	Length 623
Best Local Similarity	27.7%	Pred. No. 2.2e-36		
Matches 157	Conservative 75	Mismatches 232	Indels 103	Gaps 19
Qy	63	AETWDSQSDMEWETTVDENTKKOVFI	FD	SLVKKCLFEVLNTKNIFDGDVNVWFVHGWK 122
Db	31	AEKEWELPPDSMDLNLIE--QAPLTVAEKLQREFLVEWRVSK--APEALFFVQFEKG-	85	
Qy	123	DQGWCHCHVLJG-----GKDFSQAGQKWRRQLNVVYSRWLVTACNVOLTPAERIKL	173	
Db	86	DSYFHLHLIVETGVGKSMVGRVYSQIKK-----LVTRIYRGVEP-----	126	
Qy	174	RETAENNEWTLTYKHQK--TKKDYTKCVLFGNMHIAYFLTKKIKSTSPPRDGGYFLSS	231	
Db	127	----QLPNMFAVTKTRNGAGGKNKVVDC-----YIPNLLPK-----TQP-----	163	
Qy	232	DSGWK--TNF-----LKEGRHLVSKLYTDMDRETVETVTTTAQETKGR-----	275	
Db	164	ELQWATNMDQYTSACNLAEKRLVAQHL-----THVSQTQEQNKENQNPNSD	212	
Qy	276	---IQTKKEYSIKTLLKHKRVTSPEDMMWQPDYSIEMWAPGGENLLKNTLEICTL	332	
Db	213	APVIRKTSARYMELVGLVDRGITSEKQWIQEDQASYISFNAAASNSRQIKALDNASK	272	
Qy	333	TLARTKTAFLDILEKATSKLTNFSLPDTTRACRIFAFHGNWYKVCVCHAI	CCVLNRQGGKR 392	
Db	273	INSLTKTAPDYLGVONPPEDISS-----NRIYRILEMNGYDPOVAAASVFLGWAQKFGKR	327	



	Qy	393	NTVLFHCPASTGKSLTAQAIAQAVGVGCYNAAVNFPFNDCTNKKLIWVBEGAGNFQOV	452
	Dd	328	NTIWLFGPATTTGTINAEIAHAHPFYCGVNNWTNENPFNDCCVDKMWIHWBEGKMTAKVV	387
	Qy	453	NQFKATCSGOTIRIDOKGGSKOIEPTPVIMTITNEI TVVRIGCEBPHTHOPIRDRMLN	512
	Dd	388	ESAKILGGSKRVDRCKSSAQIDTPVLIVTSNTNNCAVIDGNSTTFHQPLQDRMPK	447
	Qy	513	IHLTHLPGDFGLVDKNWPMI CAWLXQNGYQSTMASYCAKWG--KVDPWSZ--NWAEPKV	569
	Dd	448	FELTKRLHDHFGVKVQEVKDFFRWASDHVTETVFHYVRKGGARKPAPNDADISEPKR	507
	Qy	570	PTPINLLGSARSFTTTPKSTPLSQNYA	596
	Dd	508	ACP-----SVAQPSTSDAEPV--DYA	527

RESULT 4  
US-09-807-802A-7  
; Sequence 7, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 546  
; TYPE: PRT  
; ORGANISM: AAV-1  
; US-09-807-802A-7

[illegible]

```

Qy 443 EEAGNFGQVQVQFKATCSQGTTRIDQKGKSGQIETPTVIMTNNENITVVRIGCERPEH 502
Db 378 EEGQMTAKVYESAKALGGSKVRVDQCKSSAQIDPTPVIVTSTNMCAVIDGNTTTFEH 437
Qy 503 TQPIRDMNLNIHLTHLPDGFGLVDKNEWPMICAMLVKNQYQSTWASYSACKWG----KVP 558
Db 438 QQPLQDMFKFELTRLEHDFGKYTKQEVKEFPRWAQDHVTEVAHEFVYRKGANKRPAP 497
Qy 559 DWSENWAEKVPVPTPINLLGSARSPFTT 585
Db 498 DDADK-SEPKRACP-----SVADPSTS 518

RESULT 5
US-09-807-802A-2
; Sequence 2, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Willson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-2

```

Query Match	11.9%	Score 431.5;	DB 4;	Length 623;
Best Local Similarity	27.5%;	Pred. No. 9.4e-36;		
Matches	156;	Conservative 64;	Mismatches 224;	Indels 123; Gaps 19;
Qy	63	AETTWDSQSEMEWE-----TTVDGMTKKQVFIFDSLVAKKCLEVLTNKNIFPDVNW	114	
Db	31	AEKWEPLPPSDMDLNLISQAPLTVAEKIQDFLVQWRVSKA-----PEALF	78	
Qy	115	FVQHEMGKQGWCHVLIG-----GKQFSQAQGWWR---RQLNVYWSRWLVTACN	162	
Db	79	FVQPEKESY-FHLHILVETTVGKSVLGRFLSQIRDKLVQVIRGIEPTLPNFAVT--	135	
Qy	163	VQLTPAERIKLREIABEDNEWVTLTLYYKHQTKKDYTKCVLFGNMIAYPLTKKISTSP	222	
Db	136	-----KTRNAGGGGNKVDECY-----IPVLLPK-----TQP-	163	
Qy	223	RDGYFLSSDSGWK-TNF-----LXGEGRHLVSKLYTDDMPETVETVTTTAQETK	272	
Db	164	-----ELQAWNTMEEYISACINLAERKLVAQHL-----THVSQTQSQN	203	
Qy	273	RGRIQTKKEYSI---KTLTKE-----LVHKVTPSPEDMMWQPPDSYIEMMAQPGENLL	323	
Db	204	KENLNPNSDAPVTRTSARYMELVGWLVDRGITSEKQIWQEDQASYISFNAAHSRSQI	263	
Qy	324	KNTLIECTLPLARTKTAFLDILIEKAETS KL-TNFSLPDTRACRIFAFHGWNTVKVCHAIC	382	
Db	264	KAALDNAGKIMALTKSAPDYLGVAPPADIKTN-----RIYRILELNGVEPAYAGSVFL	317	
Qy	383	CVLNRQGGKNTVLFHGPASTGKSIITQAIQAAGVNGVCYNAANVPFPNDCTNKLIVV	442	
Db	318	GWAQKRGKRTIWLFGPATVTKNTIAEAIAHAVPYGCNVNTNENFPFNDQCDVKKIVW	377	
Qy	443	BEAGNFQOQVNFKAICSGGTIRIDQKGSQIEPTVPIMTNNITVVRIGCEREPH	502	
Db	378	BEGKMTAKVYESAKALIGSKVRVDQCKSSAQIDPTFPIVITSNTNMCAVIDGNSTTPEH	437	

Qy	503	TPQIRDBMLNIHLTHHLPCDGLVDKQENPMICAWLVKNGVQSTMASYCAKNG	----	KYP	558
Db	438	QQPQLDRMFKFELTRLEHDFGKVTQSEKGFRRMAQDHVTEVAHEFYVRKGGANKRPAP	----		497
Qy	559	DWSENWAEPKVPTPINLIGSARSPETT		585	
Db	498	DDADK-SEPKRACP-----SVADPSTS		518	

```

RESULT 6
US-09-807-802A-5
; Sequence 5, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-5

```

Qy	: :	559 DWSNWAEPKVPPTPINLGSARSPFTT	585
Db	: :	498 DDADK-SEPGRACP-----SVADPSTS	518

```

RESULT 7
US-09-532-594B-8
; Sequence 8, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc.feature
; OTHER INFORMATION: AAV4 Rep protein 40
US-09-532-594B-8

```

```

RESULT 8
US-09-532-5948-9
; Sequence 9, Application US/09532594B
; Patent No. 6468524
;
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
;
; TITLE OF INVENTION: AA4V VECTOR AND USES THEREOF
;
; FILE REFERENCE: 14014.025202
;
; CURRENT APPLICATION NUMBER: US/09/532,594B
;
; CURRENT FILING DATE: 2000-03-22

```

```

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.8%; Score 430.5; DB 4; Length 399;
Best Local Similarity 35.3%; Pred. No. 5.3e-36;
Matches 110; Conservative 45; Mismatches 139; Indels 15; Gaps 5;

QY 291 LVHKRVTSPEWMMQPDSDYIEMMAQPGGKNTLKTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEQDQASYISFNAAASRSQIKAAALDNAGKIMALTAKSAPDYLVPAPP 66

QY 351 SKL-TNFSLPDTRACRIFAFHGMNKKVYCHAIACVNLNROGGKNTVLFHGPASTGKSIIA 409
Db 67 ADIKTN-----RIVRILELNGYEPAYAGSVFLGWAQKFGKNTIWLFGPATTKNTNIA 120

QY 410 QATAQAVGNVGCYNAAVNFNPDCTNKNLWVEEAGNFGQVQVQVQVQVQVQVQVQVQV 469
Db 121 EALAHAVPFYGCNVNTNENFPFNDCCVDKMWIWEEGKMTAKVVESAKAILGSKVRVDQK 180

QY 470 GKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKN 529
Db 181 CKSSAQIDPTPVITVTSNTNCAVIDGNSITTFEHOQPLQDRMFKFELTRLEHDFGKVTQ 240

QY 530 EWPICAWLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVPPTPINLLGSARSPTT 585
Db 241 EVKEFFRWAQDHVTEVAHEFYVRKGGANKRPAPDDADK-SEPKRACP-----SVADPSTS 294

RESULT 10
US-09-807-802A-9
; Sequence 9, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilison, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-9

Query Match      11.6%; Score 423.5; DB 4; Length 399;
Best Local Similarity 36.3%; Pred. No. 2.9e-35;
Matches 109; Conservative 38; Mismatches 136; Indels 17; Gaps 5;

QY 291 LVHKRVTSPEWMMQPDSDYIEMMAQPGGKNTLKTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEQDQASYISFNAAASRSQIKAAALDNAGKIMALTAKSAPDYLVPAPP 66

QY 351 SKL-TNFSLPDTRACRIFAFHGMNKKVYCHAIACVNLNROGGKNTVLFHGPASTGKSIIA 409
Db 67 ADIKTN-----RIVRILELNGYEPAYAGSVFLGWAQKFGKNTIWLFGPATTKNTNIA 120

QY 410 QATAQAVGNVGCYNAAVNFNPDCTNKNLWVEEAGNFGQVQVQVQVQVQVQVQVQVQV 469
Db 121 EALAHAVPFYGCNVNTNENFPFNDCCVDKMWIWEEGKMTAKVVESAKAILGSKVRVDQK 180

QY 470 GKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKN 529
Db 181 CKSSAQIDPTPVITVTSNTNCAVIDGNSITTFEHOQPLQDRMFKFELTRLEHDFGKVTQ 240

QY 530 EWPICAWLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVPPTPINLLGSARSPTT 585
Db 241 EVKEFFRWAQDHVTEVAHEFYVRKGGANKRPAPDDADK-SEPKRACP-----SVADPSTS 294

RESULT 11
US-09-171-461-3

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.8%; Score 430.5; DB 4; Length 399;
Best Local Similarity 35.3%; Pred. No. 5.3e-36;
Matches 110; Conservative 45; Mismatches 139; Indels 15; Gaps 5;

QY 291 LVHKRVTSPEWMMQPDSDYIEMMAQPGGKNTLKTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEQDQASYISFNAAASRSQIKAAALDNAGKIMALTAKSAPDYLVPAPP 66

QY 351 SKL-TNFSLPDTRACRIFAFHGMNKKVYCHAIACVNLNROGGKNTVLFHGPASTGKSIIA 410
Db 67 EDISS-----NRIYRILEMNGYDPOVAASVFLGWAQKFGKNTIWLFGPATTKNTNIAE 121

QY 411 AIAQAVGNVGCYNAAVNFNPDCTNKNLWVEEAGNFGQVQVQVQVQVQVQVQVQVQV 470
Db 122 AIAHAVPFYGCNVNTNENFPFNDCCVDKMWIWEEGKMTAKVVESAKAILGSKVRVDQK 181

QY 471 KGSQKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKN 530
Db 182 KSSAQIDPTPVITVTSNTNCAVIDGNSITTFEHOQPLQDRMFKFELTRLEHDFGKVTQ 241

QY 531 WPMICAWLVKNGYQSTMASYCAKMG--KVPDWSNNAEPKVPPTPINLLGSARSPTTTPK 587
Db 242 VKDFFRWASDHVTEVTHEFYVRKGGARKRPAPDDADISEPKRACP-----SVAQPSSTA 296

QY 588 STPLSONYA 596
Db 297 EAPV--DYA 303

RESULT 9
US-09-807-802A-11
; Sequence 11, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilison, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 322
; TYPE: PRT
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (943)..(944)
; OTHER INFORMATION: minor splice site
US-09-807-802A-11

Query Match      11.6%; Score 423.5; DB 4; Length 322;
Best Local Similarity 36.3%; Pred. No. 2e-35;
Matches 109; Conservative 38; Mismatches 136; Indels 17; Gaps 5;

```

```
; Sequence 3, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbaumer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-171-461-3

Query Match 4.0%; Score 144; DB 3; Length 276;
Best Local Similarity 24.1%; Pred. No. 3.7e-06;
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEWMMQPDYS--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
DB 10 TLVHALIDRGVSRQWQVDPAYQYFHRSKORGFK--VRHILRDVIRHMCWSRTLLDY 67
QY 344 ILEKAETSKLTNFSLPD---TRACRIFAFHGMNYYKVKVCHAI CCVLNRQGGKENTVLFHG 399
DB 68 MSSASTPS-----PDDVLRNPLYQLLLCNGYNPAVVGTTALIRWAGHQ-SNRTVWIRG 119
QY 400 PASTGKSIITAAIAQAVNGVCYNAANVPFNDCTNKNLIWVEEAGNFQGVNQPKAIC 459
DB 120 TPMSGAPYLAQAIAYCSPLVGSVDNRKSNPFEGCPDSLFWMDGGYVYDCCVGLVKQVF 179
QY 460 SGQTIIRDQK-----KGSQIETPTVIMTNENITVVRIGCEE-RPEHTQPIRDRMLNIH 514
DB 180 RGEHVILPPEGLRGNPCSELFRTPVIMYSQADICMTRLRSGLSAEHAVALGRDCMYLIR 239
QY 515 LTHHLPGDF---GLVDKNEWPMICAW 537
DB 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 12
US-09-970-711-3
; Sequence 3, Application US/09970711
; Patent No. 6773709
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbaumer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800001
; CURRENT APPLICATION NUMBER: US/09/970,711
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/171,461
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: PCT/EP97/01944
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-171-461-3

Query Match 4.0%; Score 144; DB 3; Length 276;
Best Local Similarity 24.1%; Pred. No. 3.7e-06;
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEWMMQPDYS--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
DB 10 TLVHALIDRGVSRQWQVDPAYQYFHRSKORGFK--VRHILRDVIRHMCWSRTLLDY 67
QY 344 ILEKAETSKLTNFSLPD---TRACRIFAFHGMNYYKVKVCHAI CCVLNRQGGKENTVLFHG 399
DB 68 MSSASTPS-----PDDVLRNPLYQLLLCNGYNPAVVGTTALIRWAGHQ-SNRTVWIRG 119
QY 400 PASTGKSIITAAIAQAVNGVCYNAANVPFNDCTNKNLIWVEEAGNFQGVNQPKAIC 459
DB 120 TPMSGAPYLAQAIAYCSPLVGSVDNRKSNPFEGCPDSLFWMDGGYVYDCCVGLVKQVF 179
QY 460 SGQTIIRDQK-----KGSQIETPTVIMTNENITVVRIGCEE-RPEHTQPIRDRMLNIH 514
DB 180 RGEHVILPPEGLRGNPCSELFRTPVIMYSQADICMTRLRSGLSAEHAVALGRDCMYLIR 239
QY 515 LTHHLPGDF---GLVDKNEWPMICAW 537
DB 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 13
US-09-300-909-17
; Sequence 17, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,909
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/083,942
; FILING DATE: 01-MAY-1998
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-300-909-17

Query Match 3.3%; Score 121.5; DB 3; Length 647;
Best Local Similarity 19.7%; Pred. No. 0.0041;
Matches 110; Conservative 78; Mismatches 172; Indels 197; Gaps 26;

QY 39 GKDIGMNSYKKELOEDELK-SLQGAETTWDOSE---DMEWETT-VDEMTKKQVIFDS 92
DB 100 GKXVGRNT-RGTLQELSLNVSSTQATQTVYSPDSGYGNEVEETAEEVET----- 149
QY 93 LVKCKLFEVLNTKNIPPGDWNMFVQHWGKDGQCHVLIIGGKDFSAQAGKWRRLQNVY 152
DB 150 -----VATNTINGDAEGEHGGSVREE-----CSSVDSAIDSENQDPK----- 185
QY 153 WSRWLVTACNVQLTPAERIKLREIAEDNEWVILLIYKHQOT-----KKDYTKC 200
DB 186 -----SPTAQIKLL-LOSNNKKAAMLT-QFKETYGLSFTDLVTRFKSKDTTC 230
```

```

QY 201 VLFONMIAIYFLTKKIKISTSPRRGGYFLSSDSGWKTNFLKEGHERHUVSK--LYTDDMRP 258
D 231 T--DWAAAFGVHPTIA-----EGPKTLINKYALYTH----- 260
QY 259 ETVETTITTAQETKRGRIQT-----KKEVSIKTTLKELVHKRVTSPEDEMMOPDSY 310
D 261 -----IQSLDTKQGVLIILMLRYTCGRKRVTVGKGLSTLLH-----VPESCMLEPPKL 309
QY 311 IEMMA-----QPGENLLKNTLE-----ICTLTAR-----TKTAFDL--ILEKAETSKLT-- 354
D 310 RSPVAALYWRVTGSIISVVTGTPWIORLTVIQHGIDSDVFDLSMDVMQWAFDNEYTDE 369
QY 355 -----NPSLPDTRACRIFAFHGN-----YKVKCHAIC-----CVL 385
D 370 SDIAFNAMLADCNNAALFKSNCOAKYVKDCATMCKHYKRAQKROMSMSQWIKFRCSK 429
QY 386 NROGG-----KENTVLFHGPASTGKSIIAQAIAQ 414
D 430 CDEGDMWRPIQVFLRYOGIFBIFLPCALKBFKGTGPKKNCICIVYGPANTGKSHFCWSLMH 489
QY 415 AV--GNVCYNAANVFPFNDCTNKNLIWBEAGN--FGQOVNPFKAICSGQTIRIDQK 471
D 490 FLOGTVISYNSHFWLEPLADAKLMLDDATGCTCHSYFDNYMRNALDGYAISLDEK 549
QY 472 GSKQIEPTVIMTNN 488
D 550 SLLQMKCPPLITSNTN 566

```

RESULT 14

```

US-09-107-532A-6086
; Sequence 6086, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES

```

```

; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...589
; SEQUENCE DESCRIPTION: SEQ ID NO: 6086:
US-09-107-532A-6086

Query Match 3.1%; Score 112; DB 4; Length 589;
Best Local Similarity 19.9%; Pred. No. 0.035;
Matches 94; Conservative 60; Mismatches 150; Indels 168; Gaps 23;

QY 162 NVOLITPAERIKLEIAEDNEWTLLT-YKHQTKKDYTKVLFGNMIAIYFLTKKIKIST 220
D 142 NIKLRNATTKAIRHYLDNDLFDIETPYLGKSPGEG-----ARDYLPVSR-- 187
QY 221 PPRDGGYFLS---SDSGWKTNFKEG--ERHLVSKLYTD-----DMRPE-----TVETTVT 266
D 188 ---HAGHFVALPOSQOLFQKLLMNAAGFDRIYQIVRCFERDEDLRGDRQPEFTQVDIETFL 244
QY 267 TAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPEDEMMOPDSYIEMMAQPGENLLKNT 326
D 245 TAEE-----IQVTEGLIAKVMKEVRGIEVTLFPPRM-----TYDEAMARYGSD----- 288
QY 327 LEICTLTARTKTAFDL--ILEKAETSKLTNFSLPDTRACRIF--APHGWNVYKVKCHAIC 383
D 289 -----KPDTRFDMELIDLSDTVKEVEF-----KVQMALENGGVKALNA-- 328
QY 384 VLNRGGKRNVLFGHPASTGKSIIAQAIAQAVGNVCYNAANVFPFNDCTNKNLIW-- 441
D 329 -----KGAADRYSRKMDQIQYVGQFGA-----KGLAWLK 359
QY 442 VEEAGNFGQOVNPFKAICSGQTIRIDQKSGKQIEPTVIM--TTNENITVVRIGCEER 499
D 360 VEEDGLKG-PIAKFMGEATEAIK-----ATDAKPGDLLMFGADKSEIVAAALGA-- 408
QY 500 PEHTQIRDMNLNHLTHLPGDFGLVDKNEMPMICAWLVNGYQSTWASYCAKWKGVDP 559
D 409 -----IRTR-----LGKELGLIDESKFNFL----- 428
QY 560 WSENMAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDALEP 611
D 429 WYTDWPFQFSEEGRYVSAHHEFTWPK-----EDIPLLAEDP 467

RESULT 15
US-09-134-001C-2897
; Sequence 2897, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2897
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2897

Query Match 3.1%; Score 111; DB 3; Length 564;
Best Local Similarity 21.2%; Pred. No. 0.041;
Matches 80; Conservative 62; Mismatches 130; Indels 106; Gaps 20;

QY 188 YKHQTKKDYTKVLFGNMIAIYFLTKKIKISTSPRRGGYFLSSDSGW-KTNFLKEGE-- 244

```

```
Db      87 YSFGIKKNPNDTVLMENMIKVPHI--KRFNSE-----LFVFKANGWQKIN-----GDEL 134
Qy      245 RHLVSKL---YTDMPETVETTTTAQETKRGRIQTKEVSIKTTLKELVHKVTSPE 301
Db      135 QGLISKMIQVLLVDYKPSL-----STLKNVVVDGLQKSTDV 169
Qy      302 WMMQPDPSYIEMMAQPGGENL-----LKNLTLEICTLT-LARTKTAFDLILEKAE--- 349
Db      170 BELVENEHYIGC-----GENNFDLNTQVVVKNISIDIFPKTFLNLSLSTNDVITDKIPPYF 224
Qy      350 ---TSKLTNFSLPDTRACRIFAFHGMVYKVCHAICCVLNRQGGKRNVTFLFHGPASTGKS 406
Db      225 KQYMLQLANYD-----DDLQVFLFOHT-AVLLTADTKYRRGLILYGGAKNGKS 271
Qy      407 IIAQAIQAQVGNVCYNAANVFPFND-----CTNKNLIWVEAGNFGQO---VNQ 454
Db      272 VYIELVKSFF-----YSKDIVSKPLNELEGRFDKESLIDKSLMASHEIGOSKIOEKIVND 326
Qy      455 FKAICSGQTIRIDQKSGSKQIE---PTPVIMTTNENITVVRIGCEERPEHTQPIRDRML 511
Db      327 FKLLSVESMHVDRKGK--TQVEVILDJLKFSTNAILNF-----PPEHAKALERRIN 377
Qy      512 NIHLTHHL-PCDFGLVDK 528
Db      378 IIPCEYTVKADTSLIDK 395
```

Search completed: January 22, 2005, 03:43:20  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:41:01 ; Search time 90.6667 Seconds  
(without alignments)  
2677.792 Million cell updates/sec

Title: US-10-069-056-10  
Perfect score: 3637  
Sequence: 1 MAGNAVSEVIGATNWLKEK.....RACFGAEPLKDPSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	458.5	12.6	610	9	US-09-792-630-21
2	458.5	12.6	610	10	US-09-953-351-21
3	458.5	12.6	610	13	US-10-080-376-21
4	458.5	12.6	610	14	US-10-082-671-27
5	458.5	12.6	610	14	US-10-097-100-21
6	458.5	12.6	610	14	US-10-023-208-21
7	458.5	12.6	610	14	US-10-375-192-7
8	455.5	12.5	626	9	US-09-792-630-29
9	455.5	12.5	626	10	US-09-953-351-29
10	455.5	12.5	626	13	US-10-080-376-29
11	455.5	12.5	626	14	US-10-082-671-35
12	455.5	12.5	626	14	US-10-097-100-29
13	455.5	12.5	626	14	US-10-023-208-29

14	455	12.5	627	9	US-09-792-630-25	Sequence 25, Appl
15	455	12.5	627	10	US-09-953-351-25	Sequence 25, Appl
16	455	12.5	627	13	US-10-080-376-25	Sequence 25, Appl
17	455	12.5	627	14	US-10-082-671-31	Sequence 31, Appl
18	455	12.5	627	14	US-10-097-100-25	Sequence 25, Appl
19	455	12.5	627	14	US-10-023-208-25	Sequence 25, Appl
20	445.5	12.2	625	14	US-10-423-704A-3	Sequence 3, Appl
21	437.5	12.0	537	10	US-09-254-747-10	Sequence 10, Appl
22	437.5	12.0	623	9	US-09-792-630-5	Sequence 5, Appl
23	437.5	12.0	623	10	US-09-953-351-5	Sequence 5, Appl
24	437.5	12.0	623	10	US-09-254-747-2	Sequence 2, Appl
25	437.5	12.0	623	10	US-09-254-747-11	Sequence 11, Appl
26	437.5	12.0	623	13	US-10-080-376-5	Sequence 5, Appl
27	437.5	12.0	623	14	US-10-082-671-11	Sequence 11, Appl
28	437.5	12.0	623	14	US-10-097-100-5	Sequence 5, Appl
29	437.5	12.0	623	14	US-10-023-208-5	Sequence 5, Appl
30	437.5	12.0	623	14	US-10-375-192-5	Sequence 5, Appl
31	436.5	12.0	627	9	US-09-792-630-27	Sequence 27, Appl
32	436.5	12.0	627	10	US-09-953-351-27	Sequence 27, Appl
33	436.5	12.0	627	13	US-10-080-376-27	Sequence 27, Appl
34	436.5	12.0	627	14	US-10-082-671-33	Sequence 33, Appl
35	436.5	12.0	627	14	US-10-097-100-27	Sequence 27, Appl
36	436.5	12.0	627	14	US-10-023-208-27	Sequence 27, Appl
37	435.5	12.0	623	9	US-09-792-630-13	Sequence 13, Appl
38	435.5	12.0	623	10	US-09-953-351-13	Sequence 13, Appl
39	435.5	12.0	623	13	US-10-080-376-13	Sequence 13, Appl
40	435.5	12.0	623	14	US-10-082-671-19	Sequence 19, Appl
41	435.5	12.0	623	14	US-10-097-100-13	Sequence 13, Appl
42	435.5	12.0	623	14	US-10-023-208-13	Sequence 13, Appl
43	435.5	12.0	623	14	US-10-375-192-2	Sequence 2, Appl
44	431.5	11.9	546	15	US-10-696-261-7	Sequence 7, Appl
45	431.5	11.9	546	15	US-10-696-282-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-792-630-21  
; Sequence 21, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-09-792-630-21

Query Match	12.6%	Score	458.5	DB	9	Length	610		
Best Local Similarity	27.4%	Pred. No.	4.9e-31						
Matches	158	Conservative	84	Mismatches	220	Indels	115	Gaps	22
QY	67	WDOSEMEVETTVDEMTKQVFI	DSLVKKCLPEVLNTRKIF	FGDVNWFVQHWGKDQGW	126				
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	35	WELPPESDLNLTVE--QPOLT	VADRIRRVFLYE--WNKFSQESKFFVQFEGSEY-F	88					
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	
QY	127	HCHVLIGGKDF-SQAQGWRRQLN	VYWSRWLTACNVQLTPAERIKLREI-AEDNEWVT	184					
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	89	HLHTLVETSGISSMVLGRY-----	VSQIRAQLV---KVVFQGLEPQINDWA	132					
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	
QY	185	LLTYKHQTKDYKTCVLFNGN	MIAYFLTKKISTSPRDRGGYFLSSDQWTKTNFLKEGE	244					
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	133	I-----TKVKGKANKVDSGYI	PAYLLPK-----VQPELOWMTNLD--	171					
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	
QY	245	RHLVSKLYTDDMRPETVETVT	TTVTTAQTKRGRITQKKEVSIKTKLKS-----LV	292					
		: : :	: : :	: : :	: : :	: : :	: : :	: : :	





Db 337 TNIAEAIAHTVPFGVCVNTNENFPFNDVCKMLIWEEGKMTNKVVESAKAILGGSKVR 396  
Qy 466 IDQKGSKSQIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525  
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHOQPLEDRMFKFELTKRLPPDFGK 456  
Qy 526 VDKNEPIMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEKVPPTPI-- 573  
Db 457 ITRQEVKDFEAW-----AKVQVPVTHEFKVPRELAGTKGAESLKRPLGD 502  
Qy 574 -----NLGARSAPT--TPKSTPLSONYA-LTPL 600  
Db 503 VTNTSYKLEKRLARLSFVPETPRSSDVTVDPAIRPL 539  
RESULT 4  
US-10-082-671-27  
; Sequence 27, Application US/10082671  
; Publication No. US20030049647A1  
; GENERAL INFORMATION:  
; APPLICANT: DAHIYAT, BASSIL  
; APPLICANT: LI, MIN  
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL  
; TITLE OF INVENTION: PROFILES  
; FILE REFERENCE: XEN/001  
; CURRENT APPLICATION NUMBER: US/10/082.671  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/270,781  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-10-082-671-27

Query Match 12.6%; Score 458.5; DB 14; Length 610;  
Best Local Similarity 27.4%; Pred. No. 4.9e-31;  
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;  
Qy 67 WDQSEDMEWETTVDKTKQVIFPDSLVKKCLFVLNKTNIFFPDGVNMFVQHEWKGQGW 126  
Db 35 WELPPESDLNLTLVE--QPOLTVDRIIRRVFLYE---WNKFSKQESKFFVQFEKSEY-F 88  
Qy 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLVACNVQLTAPARIKLR-I-AEDNEWVT 184  
Db 89 HLHTLVETSGISSMWLGRY-----VSQIRALV---KVVFQIEPQINDWA 132  
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPRGGYFLSSDSGKNTNFKLKE 244  
Db 133 I-----TKVKKGGANKVVDGSIYPAYLLPK-----VQPELOWAMTNLDE-- 171  
Qy 245 RHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKE-----LV 292  
Db 172 -YKLAALNLEERKRLVAQFLAESSQSEAAQ--REFSADPVKSKTSQKYMALVNLV 228  
Qy 293 HKRVTSPEMMQPDYSYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILEKAETSK 352  
Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY----- 279  
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNRQGGKNTVLFHGPASTGK 405  
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAAGSILYGCW---QRSFNKNTVWLYGPATGK 336  
Qy 406 SIIAQIAQAQVGNVGCYNAAANVPFNDCTNKNLIWVEEAGNFQGVNQPKAICSGQIR 465  
Db 337 TNIAEAIAHTVPFGVCVNTNENFPFNDVCKMLIWEEGKMTNKVVESAKAILGGSKVR 396  
Qy 466 IDQKGSKSQIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525  
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHOQPLEDRMFKFELTKRLPPDFGK 456

Qy 526 VDKNEPIMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEKVPPTPI-- 573  
Db 457 ITRQEVKDFEAW-----AKVQVPVTHEFKVPRELAGTKGAESLKRPLGD 502  
Qy 574 -----NLGARSAPT--TPKSTPLSONYA-LTPL 600  
Db 503 VTNTSYKLEKRLARLSFVPETPRSSDVTVDPAIRPL 539  
RESULT 5  
US-10-097-100-21  
; Sequence 21, Application US/10097100  
; Publication No. US2003008645A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION L  
; FILE REFERENCE: A-70814/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/097,100  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/953,351  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232,960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-10-097-100-21

Query Match 12.6%; Score 458.5; DB 14; Length 610;  
Best Local Similarity 27.4%; Pred. No. 4.9e-31;  
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;  
Qy 67 WDQSEDMEWETTVDKTKQVIFPDSLVKKCLFVLNKTNIFFPDGVNMFVQHEWKGQGW 126  
Db 35 WELPPESDLNLTLVE--QPOLTVDRIIRRVFLYE---WNKFSKQESKFFVQFEKSEY-F 88  
Qy 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLVACNVQLTAPARIKLR-I-AEDNEWVT 184  
Db 89 HLHTLVETSGISSMWLGRY-----VSQIRALV---KVVFQIEPQINDWA 132  
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPRGGYFLSSDSGKNTNFKLKE 244  
Db 133 I-----TKVKKGGANKVVDGSIYPAYLLPK-----VQPELOWAMTNLDE-- 171  
Qy 245 RHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKE-----LV 292  
Db 172 -YKLAALNLEERKRLVAQFLAESSQSEAAQ--REFSADPVKSKTSQKYMALVNLV 228  
Qy 293 HKRVTSPEMMQPDYSYIEMMAQPGENLLKNTLEICTLTARTKTAFLDILEKAETSK 352  
Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY----- 279  
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNRQGGKNTVLFHGPASTGK 405  
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAAGSILYGCW---QRSFNKNTVWLYGPATGK 336  
Qy 406 SIIAQIAQAQVGNVGCYNAAANVPFNDCTNKNLIWVEEAGNFQGVNQPKAICSGQIR 465  
Db 337 TNIAEAIAHTVPFGVCVNTNENFPFNDVCKMLIWEEGKMTNKVVESAKAILGGSKVR 396  
Qy 466 IDQKGSKSQIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525  
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHOQPLEDRMFKFELTKRLPPDFGK 456  
Qy 526 VDKNEPIMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEKVPPTPI-- 573

```
Db 457 ITKQEVKDFPFAW-----AKVQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502
Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTFL 600
Db 503 VTNYSKLEKRLSFPVETPRSDVTVDPAPLRPL 539

RESULT 6
US-10-023-208-21
; Sequence 21, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RTI/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-023-208-21

Query Match 12.6%; Score 458.5; DB 14; Length 610;
Best Local Similarity 27.4%; Pred. No. 4.9e-31;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

Qy 67 WQSEDMEWETTVDEMTHKQVFIQFISLVKKCLFEVLNKNIPFGDVNVFVQHEWKGQGW 126
Db 35 WELPPESDLNLTVE--QPQTVADRIRRVFLYE---WNKFSKQESKFFVQFEKSGEY-F 88
Qy 127 HCHVLIGGKDF-SQAQKGKWRRLNVYWSRWLTACNVQLTPAERIKLREI-AEDNEWVT 184
Db 89 HLHVLTVETSGISSWVLGRY-----VQIRALV---KVVFQIEPQINDWA 132
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSQKWNFLKEGE 244
Db 133 I-----TKVKKGGANKVVDGSIYAYLLPK-----VQELQWMTNLD-- 171
Qy 245 RHLVSKLYTDDMPETVETTTTAQETKGRIOPTKEVSIKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQSQBAASQ--REFSADPVIKSTOKYNALVNLV 228
Qy 293 HKRVTSPEMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSK 352
Db 229 EHGITSKQWIQENQESYLSFNSSTGNSRSQIKAAALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIFAFHGN--YV-KVCHAI CCVLNRQGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QRSFNKRNVTWLYGPATTKG 336
Qy 406 SIITAAIAQAVGNVGCYNAANVPFNDCTNKLWVEEAGNFQGVQVQKCAICSGOTIR 465
Db 337 TNIAEALAHVTVPYGCNVNTNENPFNDCCDKMLIWEESKMTNKVVEAKAILGSGKVR 396
Qy 466 IDQKGGSKQIEPTPVIMTTNENITVVRICCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQQPLEDRMPKFEFTKELPDPFGK 456
Qy 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPPI-- 573
Db 457 ITKQEVKDFPFAW-----AKVQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502

Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTFL 600
Db 503 VTNYSKLEKRLSFPVETPRSDVTVDPAPLRPL 539
```

```
RESULT 7
US-10-375-192-7
; Sequence 7, Application US/10375192
; Publication No. US20030224404A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Drittanti, Lila
; TITLE OF INVENTION: HIGH-THROUGHPUT DIRECTED EVOLUTION OF NUCLEIC ACIDS BY RATIONAL
; FILE REFERENCE: 37851-918
; CURRENT APPLICATION NUMBER: US/10/375,192
; PRIOR FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US 60/360,085
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Adeno-associated virus 5
US-10-375-192-7

Query Match 12.6%; Score 458.5; DB 14; Length 610;
Best Local Similarity 27.4%; Pred. No. 4.9e-31;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

Qy 67 WQSEDMEWETTVDEMTHKQVFIQFISLVKKCLFEVLNKNIPFGDVNVFVQHEWKGQGW 126
Db 35 WELPPESDLNLTVE--QPQTVADRIRRVFLYE---WNKFSKQESKFFVQFEKSGEY-F 88
Qy 127 HCHVLIGGKDF-SQAQKGKWRRLNVYWSRWLTACNVQLTPAERIKLREI-AEDNEWVT 184
Db 89 HLHVLTVETSGISSWVLGRY-----VQIRALV---KVVFQIEPQINDWA 132
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSQKWNFLKEGE 244
Db 133 I-----TKVKKGGANKVVDGSIYAYLLPK-----VQELQWMTNLD-- 171
Qy 245 RHLVSKLYTDDMPETVETTTTAQETKGRIOPTKEVSIKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQSQBAASQ--REFSADPVIKSTOKYNALVNLV 228
Qy 293 HKRVTSPEMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSK 352
Db 229 EHGITSKQWIQENQESYLSFNSSTGNSRSQIKAAALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIFAFHGN--YV-KVCHAI CCVLNRQGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QRSFNKRNVTWLYGPATTKG 336
Qy 406 SIITAAIAQAVGNVGCYNAANVPFNDCTNKLWVEEAGNFQGVQVQKCAICSGOTIR 465
Db 337 TNIAEALAHVTVPYGCNVNTNENPFNDCCDKMLIWEESKMTNKVVEAKAILGSGKVR 396
Qy 466 IDQKGGSKQIEPTPVIMTTNENITVVRICCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQQPLEDRMPKFEFTKELPDPFGK 456
Qy 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPPI-- 573
Db 457 ITKQEVKDFPFAW-----AKVQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502

Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTFL 600
Db 503 VTNYSKLEKRLSFPVETPRSDVTVDPAPLRPL 539

RESULT 8
US-09-792-630-29
; Sequence 29, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
```

APPLICANT: Dahiyat, Basail I.  
TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
FILE REFERENCE: A-702957/RFT/RMS/RMK  
CURRENT APPLICATION NUMBER: US/09/792,630  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 87  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 29  
LENGTH: 626  
TYPE: PRT  
ORGANISM: Muscovy duck parvovirus  
US-09-792-630-29

Query Match 12.5%; Score 455.5; DB 9; Length 626;  
Best Local Similarity 25.2%; Pred. No. 9.3e-31;  
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKKQVFI FDSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
DB 45 WEPTGIWNEHVNLPMWTLADKI-----KNIFQRMNQFNQDETDFFQLEEGS 93

QY 123 DOGMHCHVLGGKDFSOAQGKWMRRQLNVYWSRWLVACNVQLTPAERIK---LREIAED 179  
DB 94 EY-IHLHAVCPGBCRSFVLGRY-----MSQIKDSILRDVYEG 129

QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMTAYYFLTKKISTSPRRGGYFLSDSG 234  
DB 130 KQVKIPDWFSTITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179

QY 235 WKTNFKEGEHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTKLKVHK 294  
DB 180 TAAALCKQKQELLDAFQSEMNNAVQEOASTAPL-----ISNRAKNYSNLVDWLIEM 235

QY 295 RVTSPEDDMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354  
DB 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295

QY 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIQAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQVYVGVLCGWVKEFNKRNAIWLPGATTGKTINAEIAH 350

QY 415 AVGNVCYNAANVFPNDCTKNLIWVEAGNFGQVNOFKALCSGQTIIDOKGKSK 474  
DB 351 AVPFYGCNVNTNENFPNDCVCKMLIWEKGKTNKVVESAKAILGGSVAVRDQKGSV 410

QY 475 QIEPTVIMTTNENITVVRIGCSERPEHTQPIRDRMLNIHLTHLPDGLVDKNE---- 530  
DB 411 CIEPTVLIITNDMCMIVDGNSTTMEHRIPLERMFQIVLSHLEGNFGKISKKEVKEF 470

QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568  
DB 471 FKWANDNLVPVWSEFKVPTNEQTKLTE-----PVPERANEPEPPKIWAPPTREELEE 524

QY 569 -----VPTPINLLGSARSPTTKSTPLSQNYALTPLASDLEDLALPWSPTNT 617  
DB 525 LRASPELFSVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576

QY 618 PVAGTAEQNTQTBAGSKAC-QDGLSPTWSEIEEDLRACFGAPLKKDFSE 667  
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 9  
US-09-953-351-29  
Sequence 29, Application US/09953351  
Publication No. US20030036643A1  
GENERAL INFORMATION:  
APPLICANT: Li, Min  
APPLICANT: Melander, Christian  
APPLICANT: Liu, Hong-Xiang  
APPLICANT: Jin, Cheng He  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION  
FILE REFERENCE: A-70814/RFT/RMS/RMK

CURRENT APPLICATION NUMBER: US/09/953,351  
CURRENT FILING DATE: 2001-09-14  
PRIOR APPLICATION NUMBER: US 60/232,960  
PRIOR FILING DATE: 2000-09-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 29  
LENGTH: 626  
TYPE: PRT  
ORGANISM: Muscovy duck parvovirus  
US-09-953-351-29

Query Match 12.5%; Score 455.5; DB 10; Length 626;  
Best Local Similarity 25.2%; Pred. No. 9.3e-31;  
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKKQVFI FDSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
DB 45 WEPTGIWNEHVNLPMWTLADKI-----KNIFQRMNQFNQDETDFFQLEEGS 93

QY 123 DOGMHCHVLGGKDFSOAQGKWMRRQLNVYWSRWLVACNVQLTPAERIK---LREIAED 179  
DB 94 EY-IHLHAVCPGBCRSFVLGRY-----MSQIKDSILRDVYEG 129

QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMTAYYFLTKKISTSPRRGGYFLSDSG 234  
DB 130 KQVKIPDWFSTITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179

QY 235 WKTNFKEGEHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTKLKVHK 294  
DB 180 TAAALCKQKQELLDAFQSEMNNAVQEOASTAPL-----ISNRAKNYSNLVDWLIEM 235

QY 295 RVTSPEDDMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354  
DB 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295

QY 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIQAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQVYVGVLCGWVKEFNKRNAIWLPGATTGKTINAEIAH 350

QY 415 AVGNVCYNAANVFPNDCTKNLIWVEAGNFGQVNOFKALCSGQTIIDOKGKSK 474  
DB 351 AVPFYGCNVNTNENFPNDCVCKMLIWEKGKTNKVVESAKAILGGSVAVRDQKGSV 410

QY 475 QIEPTVIMTTNENITVVRIGCSERPEHTQPIRDRMLNIHLTHLPDGLVDKNE---- 530  
DB 411 CIEPTVLIITNDMCMIVDGNSTTMEHRIPLERMFQIVLSHLEGNFGKISKKEVKEF 470

QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568  
DB 471 FKWANDNLVPVWSEFKVPTNEQTKLTE-----PVPERANEPEPPKIWAPPTREELEE 524

QY 569 -----VPTPINLLGSARSPTTKSTPLSQNYALTPLASDLEDLALPWSPTNT 617  
DB 525 LRASPELFSVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576

QY 618 PVAGTAEQNTQTBAGSKAC-QDGLSPTWSEIEEDLRACFGAPLKKDFSE 667  
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 10  
US-10-080-376-29  
Sequence 29, Application US/10080376  
Publication No. US20020172968A1  
GENERAL INFORMATION:  
APPLICANT: Li, Min  
APPLICANT: Dahiyat, Basail I.  
TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
FILE REFERENCE: A-70295-2/RFT/RMS/RMK  
CURRENT APPLICATION NUMBER: US/10/080,376  
CURRENT FILING DATE: 2000-02-19  
PRIOR APPLICATION NUMBER: US 09/792,630

```
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-080-376-29

Query Match      12.5%; Score 455.5; DB 13; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;

Qy 75 WETT-----VDEMTKKQVIFPDSLVKCLFVLNKNIP-----PGDVNMFVQHEWGK 122
Db 45 WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQWNNQFNODETDFPQLEEGS 93
Qy 123 DQWHCHVLIGGKDFSOAQGWRRQLNVWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129
Qy 180 N-----EWVTLTYKHQTKDYKCVLFGNMIAYVFLTKKISTSPRDGGYFLSSDSG 234
Db 130 KQVKIPDWSITKRGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNNPLF 179
Qy 235 WKTNFKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSITKTLKELVHK 294
Db 180 TAAALCLOKQRELLDAFQSENNAVQEDQASTAAPL-----ISNRAAKVSNLVDMLEIEM 235
Qy 295 RVTSPEDDMMQPPSYIEMMAQPGENLLKNTLIEICTLTARTKTAFDLILEKAEKTSKLT 354
Db 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295
Qy 355 NFSLPDTRACRIFAFHGNVYKVCCHAI CCVLRQGGKRNVTLPFGPASTGKSIIAQIAIQ 414
Db 296 K-----NRIYQILKLNYPQVGSVLCGWVRFKRNAINWLYGPATTGKTINIAEIAH 350
Qy 415 AVGVGVCYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAI CSQOTIRIDQKGKSK 474
Db 351 AVPFYGCNVNINENFPNDCVDMKLIWEEGKMTNKVESA KAILGSAVRVDQCKGKV 410
Qy 475 QIETPTVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNE-----530
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPDSEN-----WAEPK-----568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTREELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLELALPFWSTPNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKPKKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTATONTGEAGSKAC-QDGLSPTWSEIEEDLRACFGAEPLKDFSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 11
US-10-082-671-35
; Sequence 35, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASSIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; FILE OF INVENTION: PROFILES
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082,671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270,791
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-082-671-35

Query Match      12.5%; Score 455.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;

Qy 75 WETT-----VDEMTKKQVIFPDSLVKCLFVLNKNIP-----PGDVNMFVQHEWGK 122
Db 45 WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQWNNQFNODETDFPQLEEGS 93
Qy 123 DQWHCHVLIGGKDFSOAQGWRRQLNVWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129
Qy 180 N-----EWVTLTYKHQTKDYKCVLFGNMIAYVFLTKKISTSPRDGGYFLSSDSG 234
Db 130 KQVKIPDWSITKRGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNNPLF 179
Qy 235 WKTNFKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSITKTLKELVHK 294
Db 180 TAAALCLOKQRELLDAFQSENNAVQEDQASTAAPL-----ISNRAAKVSNLVDMLEIEM 235
Qy 295 RVTSPEDDMMQPPSYIEMMAQPGENLLKNTLIEICTLTARTKTAFDLILEKAEKTSKLT 354
Db 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295
Qy 355 NFSLPDTRACRIFAFHGNVYKVCCHAI CCVLRQGGKRNVTLPFGPASTGKSIIAQIAIQ 414
Db 296 K-----NRIYQILKLNYPQVGSVLCGWVRFKRNAINWLYGPATTGKTINIAEIAH 350
Qy 415 AVGVGVCYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAI CSQOTIRIDQKGKSK 474
Db 351 AVPFYGCNVNINENFPNDCVDMKLIWEEGKMTNKVESA KAILGSAVRVDQCKGKV 410
Qy 475 QIETPTVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNE-----530
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPDSEN-----WAEPK-----568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTREELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLELALPFWSTPNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKPKKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTATONTGEAGSKAC-QDGLSPTWSEIEEDLRACFGAEPLKDFSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 12
US-10-097-100-29
; Sequence 29, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI
; FILE REFERENCE: A-70814/RET/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
```

; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-10-097-100-29

Query Match 12.5%; Score 455.5; DB 14; Length 626;  
Best Local Similarity 25.2%; Pred. No. 9.3e-31;  
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;  
Qy 75 WETT-----VDEMTKKQVIFDLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWNEHVNLPMTLADKI-----KNIFIQRWNQFNQDETDFFQLEEGS 93  
Qy 123 DOGWHCHVLGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTTPAEIRK---LREIAD 179  
Db 94 EY-IHLHVCPCGCRSFVLGRY-----MSQIKDSILRDVYEG 129  
Qy 180 N-----EWVLLTYKHKQTKDYKCVLFGNMIAYYFLTKKISTSPDRDGGYFLSSDSG 234  
Db 130 KQVKIPDWFSTTKRGQNKVTAA-----YILHYLIPKKQ-----PELQWAFTNMPLF 179  
Qy 235 WKTNFKGEBRHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTKLKELVHK 294  
Db 180 TAAALCLQKQELLDAFOSEMNNAVQEDQASTAAPL-----ISNRAKNYSNLVDMLIEM 235  
Qy 295 RVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFDLLEKAEYSKLT 354  
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295  
Qy 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIIAQIAIQ 414  
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295  
Qy 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIIAQIAIQ 414  
Db 296 K-----NRIYQLKLNYPQYVGVLCGWVKEFNKRNALWLYGPATTKTNIAEAIAH 350  
Qy 415 AVGVGYNAAVNPFPNDCTNKNLIWBEAGNFGQVNOFKAICSGQTTIRIQKKGSK 474  
Db 351 AVDFYGCVNTNENFPNDCVDRKMLIWEKGKTNKVESAKAILGSAVRVDQCKGVS 410  
Qy 475 QIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE--- 530  
Db 351 AVDFYGCVNTNENFPNDCVDRKMLIWEKGKTNKVESAKAILGSAVRVDQCKGVS 410  
Qy 475 QIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE--- 530  
Db 411 CIEPTPVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPDSEN-----WAEPK----- 568  
Db 471 FKWANDLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPMSPTNT 617  
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTAETQNTGEAGSKAC-QDQGLSPTWSEIBEDLRACFGAEPKPKDFSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 13  
US-10-023-208-29  
; Sequence 29, Application US/10023208  
; Publication No. US20030124537A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Liu, Yuan-Ching  
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES  
; FILE REFERENCE: A-70174-1/RET/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/023,208  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/256,163  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 29

; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-10-023-208-29

Query Match 12.5%; Score 455.5; DB 14; Length 626;  
Best Local Similarity 25.2%; Pred. No. 9.3e-31;  
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;  
Qy 75 WETT-----VDEMTKKQVIFDLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWNEHVNLPMTLADKI-----KNIFIQRWNQFNQDETDFFQLEEGS 93  
Qy 123 DOGWHCHVLGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTTPAEIRK---LREIAD 179  
Db 94 EY-IHLHVCPCGCRSFVLGRY-----MSQIKDSILRDVYEG 129  
Qy 180 N-----EWVLLTYKHKQTKDYKCVLFGNMIAYYFLTKKISTSPDRDGGYFLSSDSG 234  
Db 130 KQVKIPDWFSTTKRGQNKVTAA-----YILHYLIPKKQ-----PELQWAFTNMPLF 179  
Qy 235 WKTNFKGEBRHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTKLKELVHK 294  
Db 180 TAAALCLQKQELLDAFOSEMNNAVQEDQASTAAPL-----ISNRAKNYSNLVDMLIEM 235  
Qy 295 RVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFDLLEKAEYSKLT 354  
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295  
Qy 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIIAQIAIQ 414  
Db 296 K-----NRIYQLKLNYPQYVGVLCGWVKEFNKRNALWLYGPATTKTNIAEAIAH 350  
Qy 415 AVGVGYNAAVNPFPNDCTNKNLIWBEAGNFGQVNOFKAICSGQTTIRIQKKGSK 474  
Db 351 AVDFYGCVNTNENFPNDCVDRKMLIWEKGKTNKVESAKAILGSAVRVDQCKGVS 410  
Qy 475 QIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE--- 530  
Db 411 CIEPTPVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPDSEN-----WAEPK----- 568  
Db 471 FKWANDLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPMSPTNT 617  
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTAETQNTGEAGSKAC-QDQGLSPTWSEIBEDLRACFGAEPKPKDFSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 14  
US-09-792-630-25  
; Sequence 25, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahiayat, Basil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 25  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-09-792-630-25

Best Local Similarity	25.8%;	Prod. No. 1e-30;								
Matches	168;	Conservative	76; Mismatches 265; Indels 140; Gaps 23;							
Qy	75	WETT---	VDEMTKQKQFI	FDLSVKKCL	FEVLNTKNIF-----	PGDVNWFVQHEWGK	122			
Db	45	WEPTGIWN	HEHVLN	PWTLADK1-	-----	KNIFQRMNQFQDETF	FFPQEEGS	93		
Qy	123	DOGHCHVL	IGGKDF	SOAQGKWR	RQLNVYWS	RMLVTACN	VLTPAERIK--	LRETAED	179	
Db	94	EY-IHLHCCI	-----	AQGNVS	FVLGRYS-----	-----	QIKDSILRDVYEG	130		
Qy	180	N-----	BWTVLLIY	YKHQTK	KDYTKCVL	FGNMIAYFL	TKKXISTSP	RDPGGVFLSDSG	234	
Db	131	KQVKIPDWF	SITTKRG	QGNKTVTAA	-----	YILHVLIPKKQ	-----	PELQWAFNTNPLF	180	
Qy	235	WKTNFLKEG	RHLVSKLY	TDMPETVET	TVTTAQETK	GRIGTKKEVS	IKITLKELVHK	294		
Db	181	TAAALCQK	QKEQLDAP	QOESEMNAV	QVEDQAAPEL	-----	ISNRAKNYSN	LDWMLIEM	236	
Qy	295	RVTSPEDMW	MMQPD	SYIEMMAQ	PGGEGNLLKNTLIE	ICTLTLARTK	TAFDLILEKAETSKLT	354		
Db	237	GITSEKQWL	TENKESYRS	FQATSSNNR	QVKAALENARA	EMLTKTATDYL	IGKDPVLDDIT	296		
Qy	355	NFSLPDT	TRACRI	IFAFHGWNT	VYKCHAI	CCVNLNQGGK	RNTVLPHG	PASTKTSIIAQAIQ	414	
Db	297	K-----	NRITYQIL	KLNINYP	QYYVGSV	LCGVKVRKEFN	KRNAIWL	YGPAITTKTNIAEAI	351	
Qy	415	AVGNVGCY	NAANVNF	PFPDNC	TNKNLW	VEAGNFQGVNQ	FKAICSG	QOTIRIDOKGSK	474	
Db	352	AVFPYGC	VWNTNF	FPDNCV	DKMLW	BEGQNTK	KVVSAKAIL	CGSARVVDQCKG	411	
Qy	475	QIEPTP	VIMTNN	ENITVVR	IGCB	ERPEHTQPI	RDRLMLNI	HLTHLHPDFGLVDKNE	530	
Db	412	CIEPTP	VIITSD	MCMI	VDGNS	TTMEHRI	PLEERMFQI	VLSHKLGNFGKISKEV	471	
Qy	531	--W-----	PMICAWL	VKNYQ	OSTMASY	CXKWKVPD	WSEN-----	WAEPK-----	568	
Db	472	FRWANDNL	VEVWSEF	KVPNT	BEQTKLTE	-----	PVPERANE	PSEPPKIWA	PPTRESLEBI	525
Qy	569	-----	VPTPIN	LLGSAR	SPTTPK	TSPLSQ	NYALTPLAS	DLLEDLAL	EPWSTPNT	617
Db	526	LRASPEL	FASVAPLP	-----	SSPOTS	PKRKT	GRCEYQVRC	AMHSL--	DNSMNVF	577
Qy	618	PVAGTAE	TQNTWG	BAGSKAC	-QDQLS	PWTSEI	EEDLR	ACFGAEPL	KDPFSE	667
Db	578	ERANF	PEFOSLGS	--NFCNHG	-----	WYD-----	CAFCNEL	KDDWNE	613	

Search completed: January 22, 2005, 04:03:42  
Job time : 92.6667 secs

Query Match 12.5%; Score 455; DB 10; Length 627;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:18 ; Search time 22.3333 Seconds  
(without alignments)  
2895.119 Million cell updates/sec

Title: US-10-069-056-10  
Perfect score: 3637  
Sequence: 1 MAGNAYSDEVLGATNWLKSK.....RACFGAEPLKKDPSEPLNLD 672  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3623	99.6	672	1 UYPVIM	noncapsid protein
2	3556	97.8	721	1 UYPVIM	noncapsid protein
3	3326	91.4	672	1 UYPV1	noncapsid protein
4	3322	91.3	668	1 A44276	noncapsid protein
5	2647	72.8	668	1 UYPVME	noncapsid protein
6	2644	72.7	668	1 UYPVPP	noncapsid protein
7	2642	72.6	668	1 UYPVCP	noncapsid protein
8	2485	68.3	660	1 UYPVPP	noncapsid protein
9	2476	68.1	662	1 UYPVNA	noncapsid protein
10	1621.5	44.6	392	1 UYPVIF	noncapsid protein
11	550	15.1	641	2 S41439	gene NS-1 protein
12	544.5	15.0	641	2 S41434	gene NS-1 protein
13	543	14.9	620	1 UYPVAP	noncapsid protein
14	533.5	14.7	641	2 S41861	gene NS-1 protein
15	455.5	12.5	626	2 S22209	noncapsid protein
16	424.5	11.7	536	1 UYADIA	noncapsid protein
17	352	9.7	726	1 UYPV51	noncapsid protein
18	348	9.6	671	1 UYPV19	noncapsid protein
19	169	4.6	490	2 T44050	hypothetical prote
20	168.5	4.6	490	2 JQ1630	noncapsid protein
21	144	4.0	276	2 S26428	hypothetical 31.5K
22	131.5	3.6	849	1 UYPVAD	noncapsid protein
23	127	3.5	545	2 B44054	orf2 protein - Jun
24	121.5	3.3	614	1 W1W141	E1 protein - human
25	121.5	3.3	647	1 W1W139	E1 protein - human
26	120.5	3.3	825	2 G96665	protein F22C12.12
27	117.5	3.2	1008	2 T18832	probable RNA helic
28	114.5	3.1	1098	2 B70232	hypothetical prote
29	113.5	3.1	497	2 C70454	transcription regu

ALIGNMENTS

RESULT 1

UYPVIM

C;Species: minute virus of mice, murine parvovirus  
C;Date: 14-Nov-1983 #sequence\_revision 28-Aug-1985 #text\_change 09-Jul-2004  
C;Accession: A03696

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.  
Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.  
A;Reference number: A03696; MUID:83143341; PMID:6298737  
A;Accession: A03696  
A;Molecule type: DNA  
A;Residues: 1-672 <AST>

A;Cross-references: UNIPROT:P031134; EMBL:V01115  
A;Superfamily: parvovirus noncapsid protein  
C;Keywords: noncapsid protein

Query Match 99.6%; Score 3623; DB 1; Length 672;  
Best Local Similarity 99.7%; Pred. No. 9.2e-255;  
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLGATNWLKESKNOEVSFVFKENVQNLNGKDIGNSYKKLODELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKESKNOEVSFVFKENVQNLNGKDIGNSYKKLODELKSLQ	60
Qy	61	RGAEITWDQSEDMEWETTVDKTKQVIFPDSLVKKCLFEVLNTKNIFFPGDVNMFVQHEW	120
Db	61	RGAEITWDQSEDMEWETTVDKTKQVIFPDSLVKKCLFEVLNTKNIFFPGDVNMFVQHEW	120
Qy	121	GKQGWCHVLIIGKDFSQAGKWRRLQNLVYWSRWLVTAACNVQLTPAERIKLREIAEDN	180
Db	121	GKQGWCHVLIIGKDFSQAGKWRRLQNLVYWSRWLVTAACNVQLTPAERIKLREIAEDN	180
Qy	181	EWTLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPRDGGVFLSDSGWKTNFL	240
Db	181	EWTLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPRDGGVFLSDSGWKTNFL	240
Qy	241	KEGERHLVSKLYTDDMRPETVTTTAQSTKRGRIQTKEVSIKTTLKELVHKRVTSPE	300
Db	241	KEGERHLVSKLYTDDMRPETVTTTAQSTKRGRIQTKEVSIKTTLKELVHKRVTSPE	300
Qy	301	DNMMPDSDSIENMAOPGGENLLKNTLEICTLTARTKTAFDILEKAEKSTKLTNFSLPD	360
Db	301	DNMMPDSDSIENMAOPGGENLLKNTLEICTLTARTKTAFDILEKAEKSTKLTNFSLPD	360
Qy	361	TRACRIFAFHGWYVVKVCHACVCLNRQGGKRNVLPHGPASTGKSIIAQAIQAVGNVG	420
Db	361	TRTCRIFAFHGWYVVKVCHACVCLNRQGGKRNVLPHGPASTGKSIIAQAIQAVGNVG	420
Qy	421	CYNAANVNFNDCTNKNLIWVEAGNFQGVNQKAIKCSGQTIKIDQKGGKQIETPT	480
Db	421	CYNAANVNFNDCTNKNLIWVEAGNFQGVNQKAIKCSGQTIKIDQKGGKQIETPT	480

E1 protein - human  
E1 protein - human  
anfa protein - Azo  
hypothetical prote  
E1 protein - human  
probable ATP bindi  
structural polypro  
hypothetical prote  
hypothetical prote  
E1 protein - human  
proteinase ClpX [i  
dystrophin - mouse  
hypothetical WW do  
type I site-specif  
TipC protein - sli  
large T antigen -

QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNEWPMICAWLVK 540  
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNEWPMICAWLVK 540  
QY 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600  
DB 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600  
QY 601 ASDLELALPEWSTPNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEEDLRACFGAEP 660  
DB 601 ASDLELALPEWSTPNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEEDLRACFGAEP 660  
QY 661 LKXDFSEPLNLD 672  
DB 661 LKXDFSEPLNLD 672

RESULT 2  
UYPVIM  
noncapsid protein NS1 - minute virus of mice (strain MVMi)  
C:Species: minute virus of mice, murine parvovirus  
C>Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
C:Accession: A23008; A29510  
R;Sahli, R.; McMaster, G.K.; Hirt, B.  
Nucleic Acids Res. 13, 3617-3633, 1985  
A>Title: DNA sequence comparison between two tissue-specific variants of the autonomous  
A:Reference number: A23008; MUID:85242059; PMID:3855242  
A:Accession: A23008  
A:Molecule type: DNA  
A:Residues: 1-721 <SA>  
A:Cross-references: UNIPROT:Q84363; EMBL:X02481  
R;Astell, C.R.; Gardiner, E.M.; Tattersall, P.  
J. Virol. 57, 656-669, 1986  
A>Title: DNA sequence of the lymphotropic variant of minute virus of mice, MVM(1), and c  
A:Reference number: A29510; MUID:86115415; PMID:3502703  
A:Accession: A29510  
A:Molecule type: DNA  
A:Residues: 1-645, 'I', 647-721 <AST>  
A:Cross-references: EMBL:M12033; NID:g332289; PIDN:AAA69566.1; PID:g825477  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 97.8%; Score 3556; DB 1; Length 721;  
Best Local Similarity 97.6%; Pred. No. 7.4e-250;  
Matches 656; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNEVNLQNGDKDIGWNSYKKELODELKSLQ 60  
DB 50 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKTEDEVQNLNGDKDIGWNNYKKELODELKSLQ 109  
QY 61 RGAETTDQSEDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNVFVQHEW 120  
DB 110 RGAETTDQSEDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNVFVQHEW 169  
QY 121 GKDGWHCHVLIGKDFSOAQGWRRQLNLYVSRWLVTACNVQLTPAERIKLREIAEDN 180  
DB 170 GKDGWHCHVLIGKDFSOAQGWRRQLNLYVSRWLVTACNVQLTPAERIKLREIAEDS 229  
QY 181 EWTLLTYKHQTKDYTKCVLFNGMIAYFLTKKKISTSPPRDGGYFLSDSGSWKTNFL 240  
DB 230 EWTLLTYKHQTKDYTKCVLFNGMIAYFLTKKKISTSPPRDGGYFLSDSGSWKTNFL 289  
QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300  
DB 290 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 349  
QY 301 DMMWQPDSEYEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360  
DB 350 DMMWQPDSEYEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 409  
QY 361 TRACRIFAFHGWNYKVCCHAI CCVLNRQGGKRNVTLPFHGPASTGKSI IAQIAQAVGNVG 420

DB 410 TRTCKIFAFHGWNYKVCCHAI CCVLNRQGGKRNVTLPFHGPASTGKSI IAQIAQAVGNVG 469  
QY 421 CYNAAVNVPENDCTNKNLIWVEEAGNFQGVNOFKAI CSGQTIRIDQKKGSKQIEPTP 480  
DB 470 CYNAAVNVPENDCTNKNLIWVEEAGNFQGVNOFKAI CSGQTIRIDQKKGSKQIEPTP 529  
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNEWPMICAWLVK 540  
DB 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNEWPMICAWLVK 589  
QY 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600  
DB 590 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTTPKSTPLSQNYALTPL 649  
QY 601 ASDLELALPEWSTPNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEEDLRACFGAEP 660  
DB 650 ASDLELALPEWSTPNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEEDLRACFGAEP 709  
QY 661 LKXDFSEPLNLD 672  
DB 710 LKXDFSEPLNLD 721

RESULT 3  
UYPVV1  
noncapsid protein NS1 - parvovirus H1  
C:Species: parvovirus H1  
A:Note: host Homo sapiens (man)  
C>Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 09-Jul-2004  
C:Accession: A03695  
R;Rhode III, S.L.; Paradiso, P.R.  
J. Virol. 45, 173-184, 1983  
A>Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid  
A:Reference number: A03695; MUID:83112183; PMID:6823009  
A:Accession: A03695  
A:Molecule type: DNA  
A:Residues: 1-672 <RHO>  
C:Cross-references: UNIPROT:P03133; EMBL:X01457; NID:g60993; PIDN:CAA5689.1; PID:g60994  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 91.4%; Score 3326; DB 1; Length 672;  
Best Local Similarity 91.1%; Pred. No. 3.3e-233;  
Matches 612; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNEVNLQNGDKDIGWNSYKKELODELKSLQ 60  
DB 1 MAGNAYSDEVLGATNWLKDKSSQEVFSFVKNEVNLQNGDKDIGWNSYKKELODELKSLQ 60  
QY 61 RGAETTDQSEDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNVFVQHEW 120  
DB 61 RGAETTDQSEDMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI APSNVTWVQHEW 120  
QY 121 GKDGWHCHVLIGKDFSOAQGWRRQLNLYVSRWLVTACNVQLTPAERIKLREIAEDN 180  
DB 121 GKDGWHCHVLIGKDFSOAQGWRRQLNLYVSRWLVTACNVQLTPAERIKLREIAEDS 180  
QY 181 EWTLLTYKHQTKDYTKCVLFNGMIAYFLTKKKISTSPPRDGGYFLSDSGSWKTNFL 240  
DB 181 EWTLLTYKHQTKDYTKCVLFNGMIAYFLTKKKISTSPPRDGGYFLSDSGSWKTNFL 240  
QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300  
DB 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300  
QY 301 DMMWQPDSEYEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360  
DB 301 DMMWQPDSEYEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSWAS 360  
QY 361 TRACRIFAFHGWNYKVCCHAI CCVLNRQGGKRNVTLPFHGPASTGKSI IAQIAQAVGNVG 420  
DB 361 TRICRIFAEHGWNYKVCCHAI CCVLNRQGGKRNVTLPFHGPASTGKSI IAQIAQAVGNVG 420



```
QY 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPPTTKSTPLSQNYALTPL 600
DB 541 NGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPPTTKSTPLSQNYALTPL 600
QY 601 ASLDLALPWPSTPNTPVAGTAETONTGAGSKACODGQSLSTWSEIEEDLRACFGEAP 660
DB 601 ASLDLALPWPSTPNTPVAGTAETONTGAGSKACODGQSLSTWSEIEEDLRACFGEAP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LESDFNEELTLD 672

RESULT 4
A44276
noncapsid protein NS1 - parvovirus LuIII
C:Species: parvovirus LuIII
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: A44276
R:Diffot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A:title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A:Reference number: A44276; MUID:93297126; PMID:8517025
A:Accession: A44276
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-668 <DIP>
A:Cross-references: UNIPROT:P36311; GB:M81888
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.3%; Score 3322; DB 1; Length 668;
Best Local Similarity 91.4%; Pred. No. 6.3e-233;
Matches 614; Conservative 23; Mismatches 31; Indels 4; Gaps 2;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVQNGDKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLTGTTNMLKDKSNQEVFSFVKENEDVQNGKNGIGNWSYKKELODELKSLQ 60
QY 61 RGAETTDQSEDMEWETTVDKMTKKQVFIPLSLVKKCLFEVLTKNIFPGDVNMFVQHEW 120
DB 61 RGAETTDQSEDMEWESSDELTKQVFIPLSLVKKCLFEVLSTKNIAPSDVTWVQHEW 120
QY 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAEDN 180
DB 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAEDQ 180
QY 181 EWTLLTYKHQTKDYKCVCFGNMVAIYFLTKKKISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHQTKDYKCVCFGNMVAIYFLTKKKICTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTSPE 300
QY 301 DNMWMPDYSYIENMAOPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSLPD 360
DB 301 DNMWMPDYSYIENMAOPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSLPD 360
QY 361 TRACRIFAFHGMVYKVKCHAIACVNLNROGGKRNITVLFHGPASTGKSIIAQAIAQAGNVG 420
DB 361 TRACRIFAFHGMVYKVKCHAIACVNLNROGGKRNITVLFHGPASTGKSIIAQAIAQAGNVG 420
QY 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
```

```
DB 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPPTTKSTPLSQNYALTPL 600
DB 541 NGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLLGSARSPPTTKSTPLSQNYALTPL 597
QY 601 ASLDLALPWPSTPNTPVAGTAETONTGAGSKACODGQSLSTWSEIEEDLRACFGEAP 660
DB 598 -SDLDLALPWPSTPNTPVAGTAETONTGAGSKACODGQSLSTWSEIEEDLRACFGEAP 660
QY 661 LKXDFSEPLNLD 672
DB 657 WKSDSEQLNLD 668

RESULT 5
UYPVME
noncapsid protein NS1 - mink enteritis virus (strain Abaehiri)
C:Species: mink enteritis virus, MEV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinag
J. Gen. Virol. 72, 867-875, 1991
A:title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: A38350
A:Molecule type: DNA
A:Residues: 1-668 <XAR>
A:Cross-references: UNIPROT:P27438; GB:D00765; NID:g222435; PIDN:BAA00662.1; PID:g222436
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 72.8%; Score 2647; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 5.9e-184;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVQNGDKDIGNWSYKKELODELKSLQ 60
DB 1 MSGNQITEEVMEGVNMLKHAENAEAFSVFKCDNVQNGKDVHNNYTKPIQNEELTSLI 60
QY 61 RGAETTDQSDS--EDMEWETTVDKMTKKQVFIPLSLVKKCLFEVLTKNIFPGDVNMFVQH 118
DB 61 RGAETAMDQTEEBEEMDSEVDSLAKKQVQTPDALIKKCLFEVFSKNIEPNECWFQIH 120
QY 119 ENKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVATCNVQLTPAERIKLREIAE 178
DB 121 ENKQDQGHCHVLHSHKLNQQAATGKWLRRQMNYSRWLVATCNVQLTPAERIKLREIAE 180
QY 179 DNEWVTLTYKHQTKDYKCVLFGNMIAYFLTKKKISTSPPRDGGYFLSSDSGWKTN 238
DB 181 DSEWVTLTYRHQTKDYKCVNWFHGNMIAYFLTKKKI--VHMTKESGYFLSDSGWKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTS 298
DB 240 FMKYQDRHTVSTLYTSEQMKPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTS 299
QY 299 PEDWMMQPDYSYIENMAOPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSL 358
DB 300 PEDWMMQPDYSYIENMAOPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSL 359
QY 359 PDTRACRIFAFHGMVYKVKCHAIACVNLNROGGKRNITVLFHGPASTGKSIIAQAIAQAVGN 418
DB 360 ANSRTCOIFRMHGMVYKVKCHAIACVNLNROGGKRNITVLFHGPASTGKSIIAQAIAQAVGN 419
QY 419 VGCYNAANVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEP 478
DB 420 VGCYNAANVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAICSGQTIRIDQKGSQKQIEP 479
QY 479 TPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWL 538
```

Db 480 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLKVLKPGDGLVDKEWELICAWL 539  
Qy 539 VKNGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLGSARSPTTTPKSTPLSQNYALT 598  
Db 540 VKHGYSTWANYTHHWKVPEDNWAEPKIQGGI - ISPOCKOLETQAASNPQSDHVL 598  
Qy 599 PLASDLEDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658  
Db 599 PLTPDVVDLALPEWSTPNTPIAETA-NQOSNQLG-VTHKVOQASPTWSEIEADLRAIFTS 656  
Qy 659 EPLKDFSEPLN 670  
Db 657 EQLEEDFRDLD 668

## RESULT 6

UYPVFP  
noncapsid protein NS1 - feline panleukopenia virus (strain 193)  
C/Species: feline panleukopenia virus, FPLV  
C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C/Accession: A36608  
R/Martyn, J.C.; Davidson, B.E.; Studdert, M.J.  
J. Gen. Virol. 71, 2747-2753, 1990  
A/Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus  
A/Reference number: A36608; MUID:91073139; PMID:2174965  
A/Accession: A36608  
A/Molecule type: DNA  
A/Residues: 1-668 <MAR>  
A/Cross-references: UNIPROT:P24842; GB:X55115; NID:g60863; PIDN:CAA38910.1; PID:g60864  
C/Superfamily: parvovirus noncapsid protein  
C/Keywords: noncapsid protein

Query Match 72.7%; Score 2644; DB 1; Length 668;  
Best Local Similarity 73.1%; Pred. No. 9.8e-184;  
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLNGKDIGNSYKKELODELKSLQ 60  
Db 1 MSGNQYTEEVMEGVNWLKHAENAFSVFKCDNVQLNGKDVRRNNTKPIQNEELTSLI 60  
Qy 61 RGAETTWDS--EDMEWETTVDKTKOVFI FDSLVKKCLFEVLNTKNI FPGDVNMFVQH 118  
Db 61 RGAQTAMDQTEEEEMDWESEVDSLAKKQVTFDALIKKCLFEVFSKNIEPNECVWFIOH 120  
Qy 119 EWGKDOGWCHVLIGGKDFSOAGKWRRLQNVYWSRLVATCNVLTTPAERIKLREIAE 178  
Db 121 EWGKDOGWCHVLHLSKNLQATGKWLRRQNMWYWSRLVTLCSVNLTPTEKIKLREIAE 180  
Qy 179 DNEWVTLTYKHQTKDYKCVLFGNMIAIYFLTKKKISTSPRGGYFLSSDSGWKTN 238  
Db 181 DSEWVTLTYRHQTKDYKCVLVKMGVHFGNMIAIYFLTKKKI - VHMTKESGYFLSDSGWKN 239  
Qy 239 FLKEGERHLVSKLYTDDMRPETVTTTAQETKRGRIQTKKEVSIKTTLKEIVHKRVTS 298  
Db 240 FMKYQDRHTVSTLYTEQMKPETVTTTAQETKRGRIQTKKEVSIKTLRDLVSKRVTS 299  
Qy 299 PEDWMMQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKETSCLTNFSL 358  
Db 300 PEDWMMQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKADNKLTNFSL 359  
Qy 359 PDTRACRIFAFHGNVYKVCCHAI CCVLRQGGKRNVTLVFHPGASTGKSIQAIAQAVGN 418  
Db 360 ANSRCTQIFRMHGNWIKVCHAIACVLRQGGKRNVTLVFHPGASTGKSIQAIAQAVGN 419  
Qy 419 VGCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 478  
Db 420 VGCYNAANVPNDCTNKNLIWEEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 479  
Qy 479 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDKNWPMICAWL 538  
Db 480 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLKVLKPGDGLVDKEWELICAWL 539

Qy 539 VKNGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLGSARSPTTTPKSTPLSQNYALT 598  
Db 540 VKHGYSTWANYTHHWKVPEDNWAEPKIQGGI - ISPOCKOLETQAASNPQSDHVL 598  
Qy 599 PLASDLEDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658  
Db 599 PLTPDVVDLALPEWSTPNTPIAETA-NQOSNQLG-VTHKVOQASPTWSEIEADLRAIFTS 656  
Qy 659 EPLKDFSEPLN 670  
Db 657 EQLEEDFRDLD 668

## RESULT 7

UYPVCP  
noncapsid protein NS1 - canine parvovirus (strain N)  
C/Species: canine parvovirus, CPV  
C/Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C/Accession: A29962  
R/Reed, A.P.; Jones, E.V.; Miller, T.J.  
J. Virol. 62, 265-276, 1988  
A/Title: Nucleotide sequence and genome organization of canine parvovirus.  
A/Reference number: A29962; MUID:88062992; PMID:2824850  
A/Accession: A29962  
A/Molecule type: DNA  
A/Residues: 1-668 <REE>  
A/Cross-references: UNIPROT:P12929; EMBL:M19296; NID:g333438; PIDN:AAA67459.1; PID:g33344  
C/Superfamily: parvovirus noncapsid protein  
C/Keywords: noncapsid protein

Query Match 72.6%; Score 2642; DB 1; Length 668;  
Best Local Similarity 73.1%; Pred. No. 1.4e-183;  
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLNGKDIGNSYKKELODELKSLQ 60  
Db 1 MSGNQYTEEVMEGVNWLKHAENAFSVFKCDNVQLNGKDVRRNNTKPIQNEELTSLI 60  
Qy 61 RGAETTWDS--EDMEWETTVDKTKOVFI FDSLVKKCLFEVLNTKNI FPGDVNMFVQH 118  
Db 61 RGAQTAMDQTEEEEMDWESEVDSLAKKQVTFDALIKKCLFEVFSKNIEPNECVWFIOH 120  
Qy 119 EWGKDOGWCHVLIGGKDFSOAGKWRRLQNVYWSRLVATCNVLTTPAERIKLREIAE 178  
Db 121 EWGKDOGWCHVLHLSKNLQATGKWLRRQNMWYWSRLVTLCSVNLTPTEKIKLREIAE 180  
Qy 179 DNEWVTLTYKHQTKDYKCVLFGNMIAIYFLTKKKISTSPRGGYFLSSDSGWKTN 238  
Db 181 DSEWVTLTYRHQTKDYKCVLVKMGVHFGNMIAIYFLTKKKI - VHMTKESGYFLSDSGWKN 239  
Qy 239 FLKEGERHLVSKLYTDDMRPETVTTTAQETKRGRIQTKKEVSIKTTLKEIVHKRVTS 298  
Db 240 FMKYQDRHTVSTLYTEQMKPETVTTTAQETKRGRIQTKKEVSIKTLRDLVSKRVTS 299  
Qy 299 PEDWMMQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKETSCLTNFSL 358  
Db 300 PEDWMMQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKADNKLTNFSL 359  
Qy 359 PDTRACRIFAFHGNVYKVCCHAI CCVLRQGGKRNVTLVFHPGASTGKSIQAIAQAVGN 418  
Db 360 ANSRCTQIFRMHGNWIKVCHAIACVLRQGGKRNVTLVFHPGASTGKSIQAIAQAVGN 419  
Qy 419 VGCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 478  
Db 420 VGCYNAANVPNDCTNKNLIWEEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 479  
Qy 479 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDKNWPMICAWL 538  
Db 480 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLKVLKPGDGLVDKEWELICAWL 539  
Qy 539 VKNGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLGSARSPTTTPKSTPLSQNYALT 598  
Db 540 VKHGYSTWANYTHHWKVPEDNWAEPKIQGGI - ISPOCKOLETQAASNPQSDHVL 598



Db 421 YNAAVNPFNDCTNKNLWIEEAGNFQNVQFKAICSGQTIRIDQKGGSKQIEPTPV 480  
Qy 482 IMTNTNITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVKN 541  
Db 481 IMTNTNITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVKN 540  
Qy 542 GYQSTMASYCAKGVKVPDSEWNAEPKVPPTINILGARSFPTPKSTPLSQNTALPLA 601  
Db 541 GYQATMASYHHWGNVDPWSEKWEPKWQTPINPTDSQIS-TSVKTSPADNNVAATPIQ 599  
Qy 602 SDLE-DLALPWSPTNTVAGTAETQNTGEAGSKACQDQQL---SPTWSEIEEDLRACFG 657  
Db 600 EBDLALPWSPTNTVAGTAETQNTGEAGSKACQDQQL---SPTWSEIEEDLRACFG 652

## RESULT 10

UYPVIF  
noncapsid protein NS1 - feline panleukopenia virus (fragment)  
C:Species: feline panleukopenia virus, FPLV  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: A03697  
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, P.; Winston, S.; Hahn, W.  
J. Virol. 55, 574-587, 1985  
A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus  
A:Reference number: A03697; MUID:85265017; PMID:2991581  
A:Accession: A03697  
A:Molecule type: DNA  
A:Residues: 1-392 <CAR>  
A:Cross-references: UNIPROT:P06431; EMBL:M10824; NID:g333474; PIDN:AAA47160.1; PID:g333474  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 44.6%; Score 1621.5; DB 1; Length 392;  
Best Local Similarity 77.5%; Pred. No. 6.7e-110; Indels 3; Gaps 3;  
Matches 306; Conservative 33; Mismatches 53;

Qy 276 IQTKVESIKTKLKHVHRTSPEDMMWQPSYIEMMAQPGGKNTLEICTITLA 335  
Db 1 IQTKVESIKTKLKHVHRTSPEDMMWQPSYIEMMAQPGGKNTLEICTITLA 60  
Qy 336 RTKTAFLDLEKATSKLTNPSLPTACRIFAFHGNVYKVAICVNLNROGKNTV 395  
Db 61 RTKTAFLDLEKATSKLTNPSLPTACRIFAFHGNVYKVAICVNLNROGKNTV 120  
Qy 396 LFHGPASTGKSIIAQAIAQAVGNVGCYNAANVPNDCTNKNLIWVEEAGNFQGVNQF 455  
Db 121 LFHGPASTGKSIIAQAIAQAVGNVGCYNAANVPNDCTNKNLIWVEEAGNFQGVNQF 180  
Qy 456 KAICSGQTIRIDQKGGSKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNHL 515  
Db 181 KAICSGQTIRIDQKGGSKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNHL 240  
Qy 516 THLHPGDFGLVDKNWPMICAWLVKGYOSTMASYCAKGVKVPDSEWNAEPKVPPTINL 575  
Db 241 VCKLPGDFGLVDKNWPMICAWLVKGYOSTMASYCAKGVKVPDSEWNAEPKVPPTINL 300  
Qy 576 LGSARSFPTPKSTPLSQNTALPLAGDLEALPWSPTNTVAGTAETQNTGEAGSKA 635  
Db 301 PG-CKDLETOASNPQSDHVLPTLTPDNDLALPWSPTNTVAGTAETQNTGEAGSKA 357  
Qy 636 CQDQLSFTWSEIEEDLRACFGAEPLKQFSEPLN 670  
Db 358 HKDVQASFTWSEIEEDLRACFGAEPLKQFSEPLN 392

## RESULT 11

S41439  
gene NS-1 protein - Aleutian mink disease virus  
C:Species: Aleutian mink disease virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S41439  
R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.  
submitted to the EMBL Data Library, January 1994

submitted to the EMBL Data Library, January 1994  
A:Description: Sequence comparison of the non-structural genes of four different types of  
A:Reference number: S41434  
A:Accession: S41439  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-641 <GOT>  
A:Cross-references: UNIPROT:Q65017; EMBL:Z29576  
C:Genetics:  
A:Introns: 586/1  
C:Superfamily: parvovirus noncapsid protein

Query Match 15.1%; Score 550; DB 2; Length 641;  
Best Local Similarity 25.8%; Pred. No. 7.8e-32;  
Matches 178; Conservative 108; Mismatches 262; Indels 142; Gaps 25;

Qy 18 KEKSNQEVFSFVFNENVQNGKIDGNSYKKELOEDE-----LKSQORGAETWD 68  
Db 8 EQRRLQDLYVQLKKEIN--DGEQVAVLFOQKTYTDKDNKPTKATPLRTTSSDLRLAFD 64  
Qy 69 QSEDMEWETTVDSEWK-----KQVFIFDSLKKCLFEVLNATKNIFFGQDVNFWVQ 117  
Db 65 SIBE-NLTASNEQNTNGINFCKLTGLKTLILLDKHVKSHRWD--NNK-----VNLIWQ 115  
Qy 118 HENGKQDQGWCHVLIG-----GKDFSOAQGWRRQLN-----VWMSRLVLTACNVQ- 164  
Db 116 IEKGKTPQPHHCCLGLGVDFRNEPDKVQKSLG-WFMKRLNKDLALYSNH---HCDIQD 170  
Qy 165 -LTPAEIRIKRETAENENWVTLTYKHQKQKDYTKCVLFQGNMAYYFLTKKISTSPPR 223  
Db 171 IKDPEADKAKLVKVEDGPTKPYKFNKQTKDYKPNVHLRDYTFYLFNKDKINTDSM- 229  
Qy 224 DGGVFLSSDGSQWKTNFKGSRHLVSKLYTDDMRPETVETV-----TTAQ 269  
Db 230 -DGFPAAGNGIVDN-LTKKERTLRMYLDEOSSDIMANDIDWEDQDQAPKVTDQDSA 287  
Qy 270 ETKRG-----RIOTKKEV-----SIKTKLKVHRTSPEDMMWQ 307  
Db 288 TTKTGTSLIWKSCATKVTSKKEVANPVQPSKLYSAQNTLDALFNVGCTFPEDMIKQS 347  
Qy 308 DSVIEMMAQPGGKNTLEICTITLARTKTAFLDLE--KAETSKLTNFSPLDTRACR 365  
Db 348 DKYLELSLEPNGPQKINTLLHMQVKTSTMTAFDCIIRFNEEDDKPLATIKDM---- 403  
Qy 366 IFAPHGNVYKVAICVNLNROGKNTVLFHGPASTGKSIIAQAIAQAVGNVGCYNA 425  
Db 404 -----GLNQLKVVCTILTKQGGKRGCIWFGPGTGKTLASLICKATVNYGMVTS 458  
Qy 426 NVNFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGGSKQIEPTPVIMT 485  
Db 459 NPNFPNTDCGNRIIWAEBGCLGNWVEDFKAITGGGVKVDTPKQPSIKGC-VIVTS 517  
Qy 486 NENITVVRIGCEERPEHTQPIRDRMLNHL-----THHLPGDFGLVDKNWPMICAWL 538  
Db 518 NTNITKTVGCVETNAHAPLQKRMKIKRMKTNPKTKITFG-----MLKRWL 566  
Qy 539 VKNGYQSTMASYCAKGVKVP-DWSENWAEKVPPTINILGARSFPTPKSTPLSQNTAL 597  
Db 567 -----NTWDRQPIQLSHEMPELYLET-----TGNSSATATKNTGNSQ---- 605  
Qy 598 TPLASDLEALPWSPTNTVAGTAETQNT 627  
Db 606 PTTAKSAESVNTENCDDTPKRGASSVPPKQH 635

## RESULT 12

S41434  
gene NS-1 protein - Aleutian mink disease virus  
C:Species: Aleutian mink disease virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S41434  
R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.  
submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus  
A;Reference number: S41434  
A;Accession: S41434  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-641 <GOT>  
A;Cross-references: UNIPROT:Q65020; EMBL:Z29577  
C;Genetics:  
A;Introns: 586/1  
C;Superfamily: parvovirus noncapsid protein

Query Match 15.08; Score 544.5; DB 2; Length 641;  
Best Local Similarity 25.28; Pred. No. 2e-31;  
Matches 182; Conservative 110; Mismatches 260; Indels 171; Gaps 26;

Qy 18 KEKSNQEVFVFKENNVQLNGKDIGNSVKKELQEDS-----LKSILQGAETWD 68  
Db 8 EQRLQELFE-KFYTE--VADGGLAWLFOQKYTTDKDKPKTKATPLRTTSSDLRLAFD 64  
Qy 69 QSEDMEWETTVDKTKQVIFDLSVKKCLFEVLNTKNI-----FPGDVNVFQHEWKGKD 123  
Db 65 SIEE-TLKTSNQCLTNNDINFCCLTLGLKTL--VLLDKHVRSHRDANKVNFVQVEKGKT 121  
Qy 124 QGWHCHVLIG-----GKDFSOAQGWRRQLN---VYWSRWLVTAACNVOLTPAERIK 172  
Db 122 QQPHIHCCGLGYFDKDEDSKDVQKSLG-WFIKLNKOLAVIWSNHHCDIQIGQSDGRADN 180  
Qy 173 LREIAEDNEWVTLTYKHOKTKDYTKCVLFGNMIAYFLTKKISTSPRDG--GYFLS 230  
Db 181 LKTIWIEGPT-KPKYKFNKQTKDYNAFTNLDRDYLILYFNKDKIT-----QEGMDGYAS 235  
Qy 231 SDSGWKTNFLKEGERHLVSKLYTDDM-----RPETVETTVAQETKRG 274  
Db 236 GNGGIIDN-LTNKERKALRWYLDQSQGILDEDIDWEDSQSAPKVTQDTSATSKTGS 294  
Qy 275 -----RIQTKKEVSI-----KTLKELVHKRVTSPEMMQPDYSYLEMM 314  
Db 295 LVMSKCATKVTSKVEAIPVKQPSKQWTSQAQNTLDDLYMFCGFTPEDMILKMSDRYLEMS 354  
Qy 315 AQPGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNPSLPDTRACRIFAFHGW 372  
Db 355 LEFNGAKINLTLLHMNQVTSNLTAFECIKFNEEDEDKPLDITIKDM-----GL 405  
Qy 373 NYVKVCHACVCLNROGKRNVTFLHFGPASTGKSIIAQIAQAVNGVGCYNAANVPFPN 432  
Db 406 NEQHLKVLCTILTQSKGKGVWFYGGTGTLLASLICKAVVNGVMTTSNPNFPWT 465  
Qy 433 DCTNNLIWVEEAGNFGQVQVNFKAICSGQTIRIDQKSGSKQIEPTPVIMTNENITVV 492  
Db 466 DCGNRNIWAEECGNIGNYVEDFKAITGGGVKVDTKNKPQSIKGS-VIVTSNTNITKV 524  
Qy 493 RIGCEERPEHTOPIRDRMLNIHL-----THLPGDFGLVDKNWPMICAWLVKNGYQS 545  
Db 525 TVGCVETNVFAEPLQKQKIRCMKPNPKVTPTG-----MLKWTI----- 566  
Qy 546 TMASYCAKWKGVDPDWSNWAEPKVPTEINLIGSARSPTTPKSTPLSNYALTEPLASDLE 605  
Db 567 -----STWDRIIP-----IKLSHEMP 581  
Qy 606 DLALPSTNTPVAGTAETQNTCEAGSKACQDQLSPTWSEIEDEDLRACFAGBPLKDF 665  
Db 582 ELYLET-SGPNSS--SATTATKST-----GNLQPTTAETAESVNTANCDDTPKRGAS 628  
Qy 666 SEP 668  
Db 629 SVP 631

RESULT 13  
UYFVAP  
noncapsid protein NS1 - Aleutian mink disease virus (strain ADV-G)  
N;Alternate names: left-ORF protein  
C;Species: Aleutian mink disease virus

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: A36760; A35529  
R;Bloom, M.E.; Alexandersen, S.; Perryman, S.; Lechner, D.; Wolfbarger, J.B.  
J. Virol. 62, 2903-2915, 1988  
A;Title: Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus  
A;Reference number: A36760; MUID:88275062; PMID:2839709  
A;Accession: A36760  
A;Molecule type: DNA  
A;Residues: 1-620 <BL2>  
A;Cross-references: UNIPROT:P24030; EMBL:M20036  
C;Superfamily: parvovirus noncapsid protein  
C;Keywords: noncapsid protein

Query Match 14.98; Score 543; DB 1; Length 620;  
Best Local Similarity 26.68; Pred. No. 2.4e-31;  
Matches 157; Conservative 104; Mismatches 229; Indels 100; Gaps 18;

Qy 18 KEKSNQEVFVFKENNVQLNGKDIGNSVKKELQEDSLOKGAETTTDQSB-DMEWE 76  
Db 38 EQRLQOLYVQLKKEIN---DGGGVAWLFOQKYTTDKDKPKTKATPLRTTSSDLRLAFD 94  
Qy 77 TTVDEMTKKQVIFDLSVKKCLFEVLNTKNI-----PGDVNVFQHEWKGQGW 126  
Db 95 SIEENLTASNEHLTNNEINFCKLTGLTKTLLIDKHVKSHERWDSNKVNLINQIEKGTQOF 154  
Qy 127 HCHVLIG-----GKDFSOAQGWRRQLN---VYWSRWLVTAACNVQ--LTPAERIKL 173  
Db 155 HIHCLGLGYFDKDEDPKDVQKSLG-WFMKRLNKLAVIYSN---HHCDIQIDKPEDRAKN 210  
Qy 174 RETAEDNEWVTLTYKHOKTKDYTKCVLFGNMIAYFLTKKISTSPRGGVFLSSDS 233  
Db 211 LKVMIEDGPTKPKYKFNKQTKDYNAFTNLDRDYLILYFNKDKINTDSM--DGYPAAAGNG 268  
Qy 234 GWKTNFLKEGERHLVSKLYTDDMRPETVETV-----TTAETKRG----- 274  
Db 269 GIVDN-LTNKERKALRWYLDQSQSDIMDANIDWEDQDAPKVTQDTSATTKGTSLIW 327  
Qy 275 -----RIQTKKEV-----SIKTKLKVHKRVTSPEMMQPDYSYLEMMAQP 317  
Db 328 KSCATKVTSKVEANVPQPSKLYSAQSTLDALFNVCFTPEDMIILKQSKYLELSLEP 387  
Qy 318 GGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNPSLPDTRACRIFAFHGWYV 375  
Db 388 NGPQKINTLLHMNQVTSNLTAFDCLIKFNEEDEDKPLLATIKDM-----GLNEQ 438  
Qy 376 KVCHACVCLNROGKRNVTFLHFGPASTGKSIIAQIAQAVNGVGCYNAANVPFPNDCT 435  
Db 439 YLKVLCTILTQSKGKGVWFYGGTGTLLASLICKATVNGVMTTSNPNFPWTDG 498  
Qy 436 NKNLWVEEAGNFGQVQVNFKAICSGQTIRIDQKSGSKQIEPTPVIMTNENITVVRI 495  
Db 499 NRNIWAEECGNIGNYVEDFKAITGGGVKVDTKNKPQSIKGC-VIVTSNTNITKVTVG 557  
Qy 496 CEEPEHTOPIRDRMLNIHL-----THLPGDFGLVDKNWPMICAWL 538  
Db 558 CVETNAEAELPKQKIRCMKPNPKVTPTG-----MLKRWL 596

RESULT 14  
S41861  
gene NS-1 protein - Aleutian mink disease virus  
C;Species: Aleutian mink disease virus  
C;Date: 15-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C;Accession: S41861; S41436  
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.  
submitted to the EMBL Data Library, January 1994  
A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus  
A;Reference number: S41434  
A;Accession: S41861  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-641 <GOT>  
A;Cross-references: UNIPROT:Q65023; EMBL:X77085; NID:G452597; PIDN:CAA54355.1; PID:G45252

## C:Genetics:

A;Introns: 586/1

C;Superfamily: parvovirus noncapsid protein

Query Match 14.7%; Score 533.5; DB 2; Length 641;  
Best Local Similarity 25.7%; Pred. No. 1.2e-30;  
Matches 178; Conservative 106; Mismatches 250; Indels 159; Gaps 27;

Qy 18 KEKSNQEVFSFVFKNNVQLANGKIDGWSYKKEIQEDS-----LKSILQKGAETIWD 68  
Db 8 EQLRLQDIYTLQKKE---VADGEGLAWLFOQKTYTDKDNKTKATPPURITSSDLRLAFD 64  
Qy 69 QSEDM-----EWETTVDMT-----KKQVFIPLSLVKKCLFEVLNTKNIPFGDVNMFVQ 117  
Db 65 SIEETLIASNEW-LTKDBINFCKLTGLKTLVLVDKHKVSHRWA-----DKINFILWQ 115  
Qy 118 HEMGKQGWCHVLIG-----GKFSQAQGWRRQLNLYVSRWLVTACNVOLTPAER 170  
Db 116 IEKGTQHFHTHCCGLGVFDKNEPKDVQKSLG-WLKKLN---RDAAIFSNHHCQDD 170  
Qy 171 IKLREIAEDN--EWV-----TLLTVKHOTKDYTKCVLFGNMIAYFLTKKISTSPPR 223  
Db 171 IKDPEAKANNLVHIEDGTPFYKYKQTKQDYNKPTLSDYVILYLFNKDKIH-----K 226  
Qy 224 DG--GYFLSSDSGWKTNFLKEGERHLVSKLYTDDMRPETVE-----TTVTT 267  
Db 227 EGMGYAAGNGGLIDN-LTNKERKALRKMYLDQSSDIDMADIDWEDGQDAPKVTQTD 285  
Qy 268 AQETKRG-----RIQTKKEV-----SIKTLKELVHKRVTSPEWMM 305  
Db 286 SATSKTGTSLIWKSCATKVTSKKEVANPVOQPSKKLYSAQNTLDALFNVCFTPEDMIK 345  
Qy 306 QPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFLDLE--KAETSCLTNFSLPDTRA 363  
Db 346 QSDXYLESLPENGPKINTLLHNNQVKTSTMTAFCLIKFNEEDKPLATIKDM-- 403  
Qy 364 CRIFAFHGNVYKCHACCVLNROGGKRNVLPHGPASTGKSIIAQAIQAQVNGVCYN 423  
Db 404 -----GLNEOYLKVLCTILTKQGGKRGCIWFGPGTGKTLASLICKATVNYGMVT 456  
Qy 424 AANVNFENDCTNKLWVEAGNFQGVNQFKAI CSQOTIRIDOKGSKQIEPTVIM 483  
Db 457 TSNFNPFWTDCGNRIIWAEECGNLGNVEDFKAITGGGVKVDYTKNQPSIKGC-VIV 515  
Qy 484 TTENITVVRIGCERPEHTOPIDRMLNIHL-----THLLPGDFGLDKNEWPMICA 536  
Db 516 TSNITITVTVGCVETNAHAEPKORMIKTCMKTINPKTKITPG-----MLKR 564  
Qy 537 WLKNGYQSTWASYCAKWKVP-----DWSENWAEKVPFPIINLLGSARGPP-TTPKSTP 590  
Db 565 WL-----NTWDRQPIQLSHEMPELYLETTGPN-----SSATSATKTTGNSQP 606  
Qy 591 LSQNYALTPLASDLELALPWPSTNTPVAGTA 623  
Db 607 TTAETASVSTADCD-----TPKRGAS 628

## RESULT 15

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N;Alternate names: nonstructural protein NS1; NS1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S52209

R;Zadoni Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52209

A;Molecule type: DNA

A;Residues: 1-626 &lt;ZAD&gt;

A;Cross-references: UNIPROT:Q83288; EMBL:X75093; NID:g609091; PIDN:CAA52983.1; PID:g60909

A;Experimental source: strain FM

C;Genetics:

A;Gene: NS1

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 12.5%; Score 455.5; DB 2; Length 626;  
Best Local Similarity 25.2%; Pred. No. 5.4e-25;  
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;

Qy 75 WEET-----VDEMTKKQVFIPLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWNGK 122  
Db 45 WEPTGIWNEHVNLPWMTLADKI-----KNIFIQRMNQFNODETDFPQLSEGS 93  
Qy 123 DQGWCHVLJGGKDFSOAQGWRRQLNLYVSRWLVTACNVQLTPAERIK---LRETAED 179  
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVVEG 129  
Qy 180 N-----EWVTLITVKKHQTQKDYTKCVLFGNMIAYFLTKKISTSPRGGVFLSSDSG 234  
Db 130 KQVKIPDWFISITTKRGQNKVTAA-----YILHVLIPKKQ-----PELQWAFNNPLF 179  
Qy 235 WKTNFKKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKELVHK 294  
Db 180 TAAALCLQKQELLDAPQESSEMNNAVQEDQAAPL-----ISNRAAKVSNLVDMWLIEM 235  
Qy 295 RVTSPEWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLEKAETSCLT 354  
Db 236 GITSEKQWLTENKESYRSFOATSSNNRQVKAALENAPAEMLLTATDYLIKGDPVLDT 295  
Qy 355 NLSLPDTRACRIFAFHGNVYKCHACCVLNROGGKRNVLPHGPASTGKSIIAQAIQA 414  
Db 296 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKRNAIWLPGPATTKTNTIAEAIH 350  
Qy 415 AVGNVGCYNAANVFPENDCTNKLWVEAGNFQGVNQFKAI CSQOTIRIDOKGSKG 474  
Db 351 AVPFYGCNVNTNENFPENDCDVKNLIWVEEGKWNKVVESAKAILLGSASVRVDQCKGVS 410  
Qy 475 QIEPTVIMTTNENITVVRIGCERPEHTOPIDRMLNIHLTHLLPGDFGLVDKNE---- 530  
Db 411 CIEPTVITSTDMCMIVDGNSTTMEHRIPLEERMFOIVLSHLEGNFGKISKKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKVPDWSN-----WAEPK----- 568  
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELEE 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPWPSTNPT 617  
Db 525 LRASPELFAVAFLP-----SSPDTSPKRTKRGYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTAETQNTGBAGSKAC-QDGLSPFTWSEIEBEDLRACFGARPLKKDFSE 667  
Db 577 ERANFPFQSLGE--NFCNQHG-----WYD-----CAFCNELKDDMNE 612

Search completed: January 22, 2005, 03:42:01

Job time : 26.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:19:28 ; Search time 106.333 Seconds  
(without alignments)  
3636.223 Million cell updates/sec

Title: US-10-069-056-10  
Perfect score: 3637  
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKKDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3623	99.6	672	1 VNC5 MUMIV	P01134 murine minu
2	3623	99.6	721	2 Q84365	Q84365 murine minu
3	3554	97.7	721	1 VNC5 MUMIM	P07300 murine minu
4	3554	97.7	721	2 Q84363	Q84363 murine minu
5	3552	97.7	672	2 Q83429	Q83429 mouse parvo
6	3345	92.0	672	2 Q8JVI4	Q8JVI4 rat minute
7	3336	91.7	672	2 Q8JVT8	Q8JVT8 Kilham rat
8	3335	91.7	672	2 P88899	P88899 Kilham rat
9	3327	91.5	672	2 Q8JVI8	Q8JVI8 rat minute
10	3326	91.4	672	1 VNC5 PAVHH	P01133 hamster par
11	3323	91.4	672	2 Q8JVT6	Q8JVT6 rat minute
12	3322	91.3	668	1 VNC5 PAVL3	P36311 parvovirus
13	3303	90.8	665	2 O71159	O71159 Kilham rat
14	2935.5	80.7	671	2 O71157	O71157 rat parvovi
15	2655	73.0	668	2 P89513	P89513 feline panl
16	2654	73.0	668	2 P89516	P89516 feline panl
17	2653	72.9	668	2 P89512	P89512 feline panl
18	2653	72.9	668	2 P89515	P89515 feline panl
19	2653	72.9	668	2 P90449	P90449 feline panl
20	2653	72.9	668	2 P90484	P90484 feline panl
21	2652	72.9	668	2 P89514	P89514 feline panl
22	2652	72.9	668	2 P90472	P90472 feline panl
23	2647	72.8	668	1 VNC5 MEVA	P27438 mink enteri
24	2644	72.7	668	1 VNC5_FPV19	P24842 feline panl
25	2642	72.6	668	1 VNC5_PAVCN	P12929 canine parv
26	2637	72.5	668	2 Q84393	Q84393 canine parv
27	2614	71.9	668	2 Q70M74	Q70M74 canine parv
28	2614	71.9	668	2 CAE47433	CAE47433 canine pa
29	2485	68.3	660	1 VNC5_PAVPN	P18547 porcine par
30	2476	68.1	662	1 VNC5_PAVPK	P52502 porcine par
31	2472	68.0	662	2 Q6RED5	Q6RED5 porcine par

32 2472 68.0 662 2 Q6TPD8  
33 2472 68.0 662 2 AAQ90279  
34 2472 68.0 662 2 AAR91039  
35 2463 67.7 662 2 Q6PS60  
36 2463 67.7 662 2 AAS93262  
37 1933 53.1 397 2 Q993M6  
38 1821.5 44.6 392 1 VNC5\_FPV  
39 550 15.1 641 2 Q65017  
40 544.5 15.0 641 2 Q65020  
41 543 14.9 590 1 VNC5\_ADVG  
42 542 14.9 620 2 Q96607  
43 533.5 14.7 641 2 Q65023  
44 465.5 12.8 610 2 Q6JL80  
45 465.5 12.8 610 2 AAR26464

## ALIGNMENTS

RESULT 1  
VNC5 MUMIV  
ID VNC5 MUMIV STANDARD; PRT; 672 AA.  
AC P01134;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 03-JUL-2004 (Rel. 44, Last annotation update)  
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVPl).  
GN Name=NS1;  
OS Murine minute virus (Murine parvovirus).  
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
OX NCBI\_TaxID=10794;  
RN [1]  
RP MEDLINE=83143341; PubMed=6298737;  
RA Astell C.R.; Thomson M.; Merchlinsky M.; Ward D.C.;  
RT "The complete DNA sequence of minute virus of mice, an autonomous  
parvovirus.";  
RL Nucleic Acids Res. 11:999-1018(1983).  
RN [2]  
RP HOMOLOGIZATION, AND INTERACTION WITH SYNCRIP.  
RX MEDLINE=99102562; PubMed=9847309;  
RA Harris C.E.; Boden R.A.; Astell C.R.;  
RT "A novel heterogeneous nuclear ribonucleoprotein-like protein  
interacts with NS1 of the minute virus of mice.";  
RL J. Virol. 73:72-80(1999).  
CC -I- FUNCTION: Seems necessary for viral DNA replication.  
CC -I- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP.  
CC -I- DOMAIN: The N-terminus (residues 1-275) possess a negative effect  
on transactivation.  
CC -I- DOMAIN: The C-terminus (residues 543-672) possess an activation  
domain.  
CC -I- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; J02275; AAA67109.1; -;  
DR EMBL; V01115; CAA24309.1; ALT\_INIT.  
DR PIR; A03696; UVPVIM.  
DR TRANSFAC; T02375; -;  
DR InterPro; IPR001257; Parvo NS1.  
DR Pfam; PF01057; Parvo NS1; I.  
KW ATP-binding; DNA replication; Noncapsid protein;  
KW Nonstructural protein.  
FT DOMAIN 1 276 Interacts with SYNCRIP.  
FT NP\_BIND 399 406 ATP (Potential).  
SQ SEQUENCE 672 AA; 78248 MW; 50298F27662E3C1D CRC64;

```
Query Match      99.6%; Score 3623; DB 1; Length 672;
Best Local Similarity 99.7%; Pred. No. 6.6e-260;
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPKNVQNLNGDKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPKNVQNLNGDKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWQSEDMEWETTVDEMCKQVFI FDSL VKKCLFEVLNTKNI FPGDVNVFVQHEW 120
Db 61 RGAETTWQSEDMEWETTVDEMCKQVFI FDSL VKKCLFEVLNTKNI FPGDVNVFVQHEW 120
Qy 121 GKQGWCHVLIIGKGFDSQAQGWRRQLNLYVSRWLVTACNVQLTPAERIKLREIAEDN 180
Db 121 GKQGWCHVLIIGKGFDSQAQGWRRQLNLYVSRWLVTACNVQLTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQKQDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSGWKTNFL 240
Db 181 EWTLLTYKHQKQDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSGWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIOTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIOTKKEVSIKTTLKELVHKRVTSPE 300
Qy 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Qy 361 TRACRIFAFHGNVYKCHAI CCVLNRQGGKRNVTLPFHGPASTGKSI IAAIAQAVNGV 420
Db 361 TRCIRIFAFHGNVYKCHAI CCVLNRQGGKRNVTLPFHGPASTGKSI IAAIAQAVNGV 420
Qy 421 CYNAAVNVFPFNDCTNKNLIWVEAGNFQGVNQFKAICSGQTTIRIDQKKGSKQIEPTP 480
Db 421 CYNAAVNVFPFNDCTNKNLIWVEAGNFQGVNQFKAICSGQTTIRIDQKKGSKQIEPTP 480
Qy 601 ASDLEDLALPWPSTPNTFPVAGTAETQNTGAGSKACODGQLSPTWSIEEDLACFAGP 660
Db 601 ASDLEDLALPWPSTPNTFPVAGTAETQNTGAGSKACODGQLSPTWSIEEDLACFAGP 660
Qy 661 LKXDFSPLNLD 672
Db 661 LKXDFSPLNLD 672

RESULT 2
Q84365 PRELIMINARY; PRT; 721 AA.
ID Q84365
AC Q84365;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Names=NS1;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MVM;
RA MEDLINE=83143341; PubMed=6298737;
RX Astell C.R.; Thomson M.; Merchinsk M.; Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
RL parvovirus.";
Nucleic Acids Res. 11:999-1018(1983).
```



```

QY 661 LKXDFSEPLNLD 672
DB 710 LKXDFSEPLNLD 721

RESULT 3
VNC5 MUMIM
ID VNC5 MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Name=NS1;
OS Murine minute virus (strain MUM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(i), and comparison with the DNA sequence of the fibrotropic
RT prototype strain."
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice."
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP (By
CC similarity).
CC -1- DOMAIN: The N-terminus (residues 1-275) possesses a negative effect
CC on transactivation (By similarity).
CC -1- DOMAIN: The C-terminus (residues 543-672) possesses an activation
CC domain (By similarity).
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02481; -; NOT ANNOTATED CDS.
DR EMBL; M12032; AAA69567.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT DOMAIN 1 276 Interacts with SYNCRIP (By similarity).
FT NP BIND 399 406 ATP (Potential).
FT FT CONFLICT 597 597 I -> L (in Ref. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 97.7%; Score 3554; DB 1; Length 672;
Best Local Similarity 97.5%; Pred. No. 8.7e-255;
Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNMLKESNEVEFSFVKENNVQLNGKDIGWNSYKKELODELSLQ 60
DB 1 MAGNAYSDEVILGATNMLKESNEVEFSFVKEDVQLNGKDIGWNNYKKELODELSLQ 60
QY 61 RGAETTWQSDMEWETTVDEMTHKQVFIQFDSLVKKCLFEVLNKNIFPGDVNWFVQHEW 120
DB 61 RGAETTWQSDMEWETTVDEMTHKQVFIQFDSLVKKCLFEVLNKNIFADVTWFWQHEW 120
QY 121 GKDQGWCHVLIIGKQDFSOAGKWRRLQNLVYWSRWLVLTACNVQLTPAERIKLREIAEDN 180

```

Query Match 97.7%; Score 3554; DB 2; Length 721;  
 Best Local Similarity 97.5%; Pred. No. 9.6e-255;  
 Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLTGATNWLKESKNOEVPFSVFKNENVOLNKGOTGNWSYKKELOEDELKSIO	60
Db	50	MAGNAYSDEVLTGATNWLKESKNOEVPFSVFKETEDVOLNKGIDGNWYKKELOEDELKSIO	109
Qy	61	RGAEITWQSDMEWETTVDEMTKKQVFIYFDSLIVKKCLFEVLNTKNIIPGDDVNVWFQHEW	120
Db	110	RGAEITWQSDMEWETTVDEMTKKQVFIYDSLIVKKCLFEVLNTKNIAPADVTWVFQHEW	169
Qy	121	GKQOQWHCHVLIGGKDFSAQOKWRRQNLNTVWSRWLVTAQCNVQLTPAERIILKLEIAEDN	180
Db	170	GKQOQWHCHVLIGGKDFSAQOKWRRQNLNTVWSRWLVTAQCNVQLTPAERIILKLEIAEDS	229
Qy	181	EWVILLTVKHKOTKDYTKCVLFQGNMIAFYFLTKKKISTSPRDCGGYFLSSDGWKTNPL	240
Db	230	EWVILLTVKHKOTKDYTKCVLFQGNMIAFYFLTKKKISTSPRDCGGYFLSSDGWKTNPL	289
Qy	241	KEGERHLVSKUYTDDMRPETVETTTTAQETKRGRIQTKKEYSIKTTLKELVHKRVTSPE	300
Db	290	KEGERHLVSKUYTDDMRPETVETTTTAQETKRGRIQTKKEYSIKTTLKELVHKRVTSPE	349
Qy	301	DWMQMPDSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAFTSKLTNFSLPD	360
Db	350	DWMQMPDSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAFTSKLTNFSLPD	409
Qy	361	TRACRIFAFHGMNYYKVCCHAICCVLNROGKRNTVLFHGPASTGKSIIAQAIAQAVGNVG	420
Db	410	TRTCKIFAFHGMNYYKVCCHAICCVLNROGKRNTVLFHGPASTGKSIIAQAIAQAVGNVG	469
Qy	421	CYNAANVPFPDNCCTKNKLIWVEEAGNFGQVNOQKAI CSQTIRIDOKGKSGKQIEBTP	480
Db	470	CYNAANVPFPDNCCTKNKLIWVEEAGNFGQVNOQKAI CSQTIRIDOKGKSGKQIEBTP	529
Qy	481	VIMTTNENITVVRIGCEBERPEHTOPIRDRLMNIHLTHLPGDFGLVDKNEWPMICAWLVK	540
Db	530	VIMTTNENITVVRIGCEBERPEHTOPIRDRLMNIHLTHLPGDFGLVDKNEWPMICAWLVK	589
Qy	541	NGYOSTWASICYAKWKVPDWSENWAEKVPPIINLLGSARBPFTTPKSTPLSONYAI TPL	600
Db	590	NGYOSTWASICYAKWKVPDWSENWAEKVPPIINSLGSARBPFTTPKSTPLSONYAI TPL	649
Qy	601	ASDLEDLALPWSPTNPVACTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFCGAEP	660
Db	650	ASDLEDLALPWSPTNPVACTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFCGAEP	709
Qy	661	LKRDFSBPLNLD 672	
Db	710	LKRDFSBPLNLD 721	
RESULT 5			
Q83429			
ID	Q83429	PRELIMINARY;	PRT; 672 AA.
AC	Q83429		
DT	01-NOV-1996	(TrEMBLrel. 01, Created)	
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	

## RESULT 5

Q83429	PRELIMINARY;	PRT;	672 AA.
ID	Q83429		
AC	Q83429;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Nonstructural protein 1.		
GN	Name=NS1;		
OS	Mouse parvovirus 1.		
OC	Viruses; ssRNA viruses; Parvoviridae; Parvovirinae; Parvovirus.		
OX	NCBI_TaxID=35340;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RF	MEDLINE=94365951; PubMed=8083985;		
RA	Ball-Goodrich L.J., Johnson E.;		
RT	"Molecular characterization of a newly recognized mouse parvovirus."		
RL	J. Virol. 68:6476-6486(1994).		
DR	EMBL; U12468; AAA61405.1; --		
DR	GO; GO:0019012; C:virion; IEA.		
DR	GO; GO:0019079; P:viral genome replication; IEA.		
DR	InterPro; IPR001257; Parvo_NS1.		
DR	PFAM; PF01057; Parvo_NS1; 1.		
KW	Nonstructural protein.		

```
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP Nonstructural protein 1.
RA Name:nsi;
RA Kilham rat virus.
RA Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
RL NCBI_TaxID=12441;
DR [1]
DR SEQUENCE FROM N.A.
DR MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses."
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.H., Soderlund-Venermo M., Pintel D., Riley L.K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF321230; AM93272.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 92.0%; Score 3345; DB 2; Length 672;
Best Local Similarity 91.7%; Pred. No. 2.8e-239;
Matches 616; Conservative 22; Mismatches 34; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESQVFSFVKENNVQNGKDIGNSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKSSQVFSFVKENNVQNGKDIGNSYKKELODELKSLQ 60
QY 61 RGAETTWQSEDMEWETTVDEMCKQVIFDLSLVKKCLFEVLNTKNIFPGDVNMFVQHEW 120
DB 61 RGAETTWQSEDMEWESAVIDMTKKQVIFDLSLVKKCLFEVLSTKNIAPSDVTWVQHEW 120
QY 121 GKQGGHCHVLIGSKDFSOAQGWRRQLNYYRSRLVTCNVQLTPAERIKLREIAEDN 180
DB 121 GKQGGHCHVLIGSKDFSOAQGWRRQLNYYRSRLVTCNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLSKKICTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQAEGRIQTKKEVSIKTKLKVHKKVTSPE 300
DB 241 KEGERHLVSKLYTDEMKEPETVETVTTAQAEGRIQTKKEVSIKTKLKVHKKVTSPE 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360
DB 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360
QY 361 TRACRIPAFHGMNKKVCHAIACVLRQGGKRNVLPHGPASTGKSIIAQAIAQGVNNG 420
DB 361 TRICRIPAEHGMNKKVCHAIACVLRQGGKRNVLPHGPASTGKSIIAQAIAQGVNNG 420
QY 421 CYNAANVNFNDCTNKNLIWEEAGNFGQVNOFKAICSGQTIRIDQKKGSKQIEPTP 480
DB 421 CYNAANVNFNDCTNKNLIWEEAGNFGQVNOFKAICSGQTIRIDQKKGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTWASYCAKGVKVPDMSNWAEPKVPPTNLLGSARSPTTTPKSTPLSONYALTPL 600
DB 541 NGYQSTWASYCAKGVKVPDMSNWAEPKVPPTNLLGSARSPTTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLACFCAEP 660
DB 601 ASDLADLALPWSPTNTPVAGTAASQNTGEGTAQCGAQRSPTWSEIEDLACFCAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 7
Q8JV28 PRELIMINARY; PRT; 672 AA.
ID Q8JV28
AC Q8JV28
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```
Db 661 LKDFSDSLTLD 672
:||||:|
661 LEKDFSDSLTLD 672

RESULT 8
P88899 PRELIMINARY; PRT; 672 AA.
AC P88899;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Non-capsid protein.
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown D.W., Like A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U79033; AAB38326.1; -.
DR GO; GO:0019079; P:Viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; I.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match 91.7%; Score 3335; DB 2; Length 672;
Best Local Similarity 91.4%; Pred. No. 1.5e-238;
Matches 614; Conservative 22; Mismatches 36; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKSSQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60

Qy 61 RGAETTTWQSDSEMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Db 61 RGAETTTWQSDSEMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120

Qy 121 GKQDQWHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDN 180
Db 121 GKQDQWHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDS 180

Qy 181 EWTLLTYKHQTKKDYTKCVLFQNMIA YFLTKKKISTSPRGGYFLSSD SGWKTNFL 240
Db 181 EWTLLTYKHHTKKDYTKCVLFQNMIA YFLSKKICTSPRGGYFLSSD SGWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQT KGRIO TKKEVSI KTKL KLVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDEMKEPETVETTTAQT AKGRIO TKKEVSI KTKL KLVHKRVTSPE 300

Qy 301 DMMWQPDYSIEMMAQPGGENLLKNTLEI CTLLARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMWQPDYSIEMMAQPGGENLLKNTLEI CTLLARTKTAFDLILEKAETSKLANFSMA 360

Qy 361 TRACRIPAFHGNVYKVCCHAI CCVLNRQGGKRN TVLFHGPASTGKSI IAAIAQAVGNVG 420
Db 361 TRTCRIPAEHGNVYKVCCHAI CCVLNRQGGKRN TVLFHGPASTGKSI IAAIAQAVGNVG 420

Qy 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGSGKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGSGKQIEPTP 480

Qy 601 ASLDELALPEPSTPNTPVAGTAETONTG EAGSKACQDGLSPTWSEIEEDLRACFGAEP 660
Db 601 ASLADLALPEPSTPNTPVAGTAETONTG EAGSKACQDGLSPTWSEIEEDLRACFSQEQ 660

Qy 661 LKDFSDSLTLD 672
:||||:|
661 LEKDFSDSLTLD 672

RESULT 9
Q8JV18 PRELIMINARY; PRT; 672 AA.
AC Q8JV18;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NS1;
OS Rat minute virus 1a.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172385;
RN [1]
RP SEQUENCE FROM N.A.
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332882; AAM93275.1; -.
DR GO; GO:0019012; C:Viralon; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; I.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76059 MW; 63D8B9B9F99E07B3 CRC64;

Query Match 91.5%; Score 3327; DB 2; Length 672;
Best Local Similarity 91.2%; Pred. No. 6e-239;
Matches 613; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKDKSQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60

Qy 61 RGAETTTWQSDSEMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Db 61 RGAETTTWQSDSEMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120

Qy 121 GKQDQWHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDN 180
Db 121 GKQDQWHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDS 180

Qy 181 EWTLLTYKHQTKKDYTKCVLFQNMIA YFLTKKKISTSPRGGYFLSSD SGWKTNFL 240
Db 181 EWTLLTYKHHTKKDYTKCVLFQNMIA YFLSKKICTSPRGGYFLSSD SGWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQT KGRIO TKKEVSI KTKL KLVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDEMKEPETVETTTAQT AKGRIO TKKEVSI KTKL KLVHKRVTSPE 300

Qy 301 DMMWQPDYSIEMMAQPGGENLLKNTLEI CTLLARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMWQPDYSIEMMAQPGGENLLKNTLEI CTLLARTKTAFDLILEKAETSKLANFSMA 360

Qy 361 TRACRIPAFHGNVYKVCCHAI CCVLNRQGGKRN TVLFHGPASTGKSI IAAIAQAVGNVG 420
Db 361 TRTCRIPAEHGNVYKVCCHAI CCVLNRQGGKRN TVLFHGPASTGKSI IAAIAQAVGNVG 420

Qy 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGSGKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGSGKQIEPTP 480
```

```
QY 481 VIMTTNITVVRIGCEERPEHTPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNITVVRIGCEERPEHTPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYSTMASYCAKWKVPDSEWNAEPKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
Db 541 NGYSTMACYCAKWKVPDSEWNAEPKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
QY 601 ASDLELALPWPSTPNTPVAGTAETONTGEGSKACQDQGLSTWSEIEEDLRACFGAEP 660
Db 601 ASDLADLALPWPSTPNTPVAGTAASQNTGEGSPTWSEIEADLRACFSQEQ 660
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 10
VNC5_PAVHH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1)..
GN Name=NS1;
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184(1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01457; CAA25689.1; -.
DR PIR; A03695; UYPPV1.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT NP BIND 399 406 ATP (Potential).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 91.4%; Score 3326; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 7.2e-238;
Matches 612; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKSEKNSQEVFVFKNENVLQNGKDIGWNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVILGATNWLKSSQEVFVFKNENVLQNGKDIGWNSYKKELODELKSLQ 60
QY 61 RGAETTDQSDMEDMEWETTVDEMVKQVFIQSLVKKCLFEVLNTKNTFPGDVNWFVQHEW 120
Db 61 RGAETTDQSDMEDMEWESAVDDMTKQVFIQSLVKKCLFEVLNTKNTAPSNVTFVQHEW 120
QY 121 GKQOGWCHVLIGKDFSQOGKWRRLQNVYWSRWLVTAQNVOLTPAERIKLREIAEDN 180
Db 121 GKQPGWCHVLIGKDFSQOGKWRRLQNVYWSRWLVTAQNVOLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLSKKXICTSPPRDGGYFLSSDSGWKTNFL 240
```

Db 1 MAGNAYSDEVILGATNWLKDKSSQEVFSFVFNKNVQNLGDKDIGNWSYKELQDELKSLQ 60  
Qy 61 RGAETTTDQSDMEWETTVDMTKKQVFIQFDSLKKCLFEVLNTKNIFFPGDVNMFVQHEW 120  
Db 61 RGAETTTDQSDMEWESAVDDMTKKQVFIQFDSLKKCLFEVLNTKNIAPSDDVTWVQHEW 120  
Qy 121 GKQGWCHVILIGKQFSQAQGWRRQLNLYVWSRWLVATCNVLTAPAEIKLREIAEDN 180  
Db 121 GKQGWCHVILIGKQFSQAQGWRRQLNLYVWSRWLVATCNVLTAPAEIKLREIAEDS 180  
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYELTKKISTSPRGGYFLSSDSGKTNFL 240  
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYELTKKISTSPRGGYFLSSDSGKTNFL 240  
Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQTGKRIOTKKEVSIKTLKELVHKRVTSPE 300  
Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTGKRIOTKKEVSIKTLKELVHKRVTSPE 300  
Qy 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNFSLPD 360  
Db 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNFSLPAN 360  
Qy 361 TRACRIFAFHGMNLYKCHAIKCVLNROGGKNTVLFHGPASTGKSIIAQAIAGVGNVG 420  
Db 361 TRTCRIFAFHGMNLYKCHAIKCVLNROGGKNTVLFHGPASTGKSIIAQAIAGVGNVG 420  
Qy 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNOFKAIKCSGQTRIDQKKGSKQIEPTP 480  
Db 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNOFKAIKCSGQTRIDQKKGSKQIEPTP 480  
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540  
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540  
Qy 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLGSARSPTTPKSTPLSONYALTPL 600  
Db 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLGSARSPTTPKSTPLSONYALTPL 600  
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIIBDLRACFGAEP 660  
Db 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIIBDLRACFGAEP 660  
Qy 661 LKQDFSEPLNLD 672  
Db 661 LKQDFSDSLTLD 672

## RESULT 12

VNCS\_PAVL3 STANDARD; PRT; 668 AA.  
AC P36311;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 05-JUN-2004 (Rel. 44, Last annotation update)  
DE Noncapsid protein NS-1 (Nonstructural protein NS1).  
GN Name=NS1;  
OS Parvovirus LuIII.  
OC Parvoviruses; Parvoviridae; Parvovirinae; Parvovirus.  
OX NCBI\_TaxID=35339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93297126; PubMed=8517025;  
RA Difford N., Chen K.C., Bates R.C., Lederma M.;  
RT "The complete nucleotide sequence of parvovirus LuIII and localization  
of a unique sequence possibly responsible for its encapsidation  
pattern.";  
RL Virology 192:339-345 (1993).  
CC -!- FUNCTION: Seems necessary for viral DNA replication.  
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M81888; -, NOT\_ANNOTATED\_CDS.  
DR PIR; A44276; A44276.  
DR InterPro; IPR001257; Parvo\_NS1.  
DR Pfam; PF01057; Parvo\_NS1; I.  
KW ATP-binding; DNA replication; Noncapsid protein;  
KW Nonstructural protein.  
FT NP\_BIND 399 406 ATP (Potential).  
SQ SSQUENCE 668 AA; 75846 MW; CAB69049F8F86B53 CRC64;

Query Match 91.3%; Score 3322; DB 1; Length 668;  
Best Local Similarity 91.4%; Pred. No. 1.4e-237;  
Matches 614; Conservative 23; Mismatches 31; Indels 4; Gaps 2;

Qy 1 MAGNAYSDEVILGATNWLKDKSSQEVFSFVFNKNVQNLGDKDIGNWSYKELQDELKSLQ 60  
Db 1 MAGNAYSDEVILGATNWLKDKSSQEVFSFVFNKNVQNLGDKDIGNWSYKELQDELKSLQ 60  
Qy 61 RGAETTTDQSDMEWETTVDMTKKQVFIQFDSLKKCLFEVLNTKNIFFPGDVNMFVQHEW 120  
Db 61 RGAETTTDQSDMEWESAVDDMTKKQVFIQFDSLKKCLFEVLNTKNIAPSDDVTWVQHEW 120  
Qy 121 GKQGWCHVILIGKQFSQAQGWRRQLNLYVWSRWLVATCNVLTAPAEIKLREIAEDN 180  
Db 121 GKQGWCHVILIGKQFSQAQGWRRQLNLYVWSRWLVATCNVLTAPAEIKLREIAEDQ 180  
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYELTKKISTSPRGGYFLSSDSGKTNFL 240  
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYELTKKISTSPRGGYFLSSDSGKTNFL 240  
Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQTGKRIOTKKEVSIKTLKELVHKRVTSPE 300  
Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTGKRIOTKKEVSIKTLKELVHKRVTSPE 300  
Qy 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNFSLPD 360  
Db 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNFSLPD 360  
Qy 361 TRACRIFAFHGMNLYKCHAIKCVLNROGGKNTVLFHGPASTGKSIIAQAIAGVGNVG 420  
Db 361 TRTCRIFAFHGMNLYKCHAIKCVLNROGGKNTVLFHGPASTGKSIIAQAIAGVGNVG 420  
Qy 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNOFKAIKCSGQTRIDQKKGSKQIEPTP 480  
Db 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNOFKAIKCSGQTRIDQKKGSKQIEPTP 480  
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540  
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540  
Qy 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLGSARSPTTPKSTPLSONYALTPL 600  
Db 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLGSARSPTTPKSTPLSONYALTPL 600  
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIIBDLRACFGAEP 660  
Db 598 -SDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIIBDLRACFGAEP 656  
Qy 661 LKQDFSEPLNLD 672  
Db 657 WKSDDSEQLNLD 668

## RESULT 13

O71159  
ID O71159 PRELIMINARY; PRT; 665 AA.  
AC O71159;  
DT 01-AUG-1998 (T-EMBLrel. 07, Created)  
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)

```
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein (fragment).
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-Mass;
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Patuzzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup.";
RL J. Virol. 72:3289-3299(1998).
DR EMBL; AF036711; AAC40695.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR Nonstructural protein.
KW NON TER
FT
SQ SEQUENCE 665 AA; 75375 MW; 778E29043417E409 CRC64;

Query Match 90.8%; Score 3303; DB 2; Length 665;
Best Local Similarity 91.4%; Pred. No. 3.6e-236;
Matches 608; Conservative 22; Mismatches 35; Indels 0; Gaps 0;

Qy 8 DEVLGATNLKEKSNQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQGAETTW 67
Db 1 DEVLGATNLKDKSSQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQGAETTW 60

Qy 68 DQSEDMEWETTVDKTKQVIFDLSVKKCLFEVLNTKNIFFPGVNNVFWQHEWGDQGW 127
Db 61 DQSEDMEWESAVDMTKQVIFDLSVKKCLFEVLNTKNIAPSDVTFVQHEWGDQGW 120

Qy 128 CHVLIGKDFSOAGKQWRRQLNYWSRWLTACNVQLTTPAERIKLREIADNFWTLLT 187
Db 121 CHVLIGKDFSQPGQKWRRLNYWSRWLTACNVQLTTPAERIKLREIADNFWTLLT 180

Qy 188 YKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFLKEGRRHL 247
Db 181 YKHQTKDYTKCVLFGNMIAYFLSKKKICTSPRGGYFLSSDSGKTNFLKEGRRHL 240

Qy 248 VSKLYTDMRPETVETTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTSPEDDMMWQP 307
Db 241 VSKLYTDMRPETVETTTTAQEAQRGRIQTKKEVSIKTLKELVHKRVTSPEDDMMWQP 300

Qy 308 DSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPDTRACRIF 367
Db 301 DSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFFMASTRICRIF 360

Qy 368 AFHGWNVYKCHACCVLNROGGKRNVLPHGPASTGKSIITAOAIAQAVGNVCYNAANV 427
Db 361 AEHGWNVYKCHACCVLNROGGKRNVLPHGPASTGKSIITAOAIAQAVGNVCYNAANV 420

Qy 428 NFPPNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTIIRIDQKGSQIEPTPIVMTTNE 487
Db 421 NFPPNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTIIRIDQKGSQIEPTPIVMTTNE 480

Qy 488 NITVVRIGCEBERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAMLVKNGYSTM 547
Db 481 NITVVRIGCEBERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPLICAMLVKNGYSTM 540

Qy 548 ASYCAKWKGVDPDMSENWAEKVPPIINLLGSARSPFTTPKSTPLSQNYALTPLASDL 607
Db 541 ACYCAKWKGVDPDMSENWAEKVPPIINLLGSARSPFTTPKSTPLSQNYALTPLASDL 600

Qy 608 ALEPWSPTNTVAGTAETQNTGEAGSKACQDQGLSPTWSETEEDLRACFGAEPLKDFSE 667
Db 601 ALEPWSPTNTVAGTAETQNTGEAGSKACQDQGLSPTWSETEEDLRACFGAEPLKDFSE 660

Qy 668 PLNLD 672
```

```
Db 661 ETLTD 665

RESULT 14
O71157 PRELIMINARY; PRT; 671 AA.
ID O71157
AC O71157
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NS1;
OS Rat parvovirus 1a.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=74581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Patuzzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup.";
RL J. Virol. 72:3289-3299(1998).
DR EMBL; AF036711; AAC40695.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; 1.
DR Nonstructural protein.
KW NON TER
SQ SEQUENCE 671 AA; 75752 MW; 9BCB39A39298D4DE CRC64;

Query Match 80.7%; Score 2935.5; DB 2; Length 671;
Best Local Similarity 81.1%; Pred. No. 6.8e-209;
Matches 545; Conservative 42; Mismatches 84; Indels 1; Gaps 1;

Qy 1 MAGNAYSEVIGATNLKEKSNQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQ 60
Db 1 MAGNAYSEVLDGVNWLKQSDTEAFSVFCKENVQNLGKDIGNSYKKELOBELKSLQ 60

Qy 61 RGAETTDQSEDMEWETTVDKTKQVIFDLSVKKCLFEVLNTKNIFFPGVNNVFWQHEW 120
Db 61 RGAETAMDQAEDEWESLQDLTKQVIFDLSVKKCLFEVLNTKNIFFPGVNNVFWQHEW 120

Qy 121 GKDQGHCHVLIGKDFSOAGKQWRRQLNYWSRWLTACNVQLTTPAERIKLREIADN 180
Db 121 GKDQGHCHVLIGKDFSOAGKQWRRQLNYWSRWLTACNVQLTTPAERIKLREIADN 180

Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFL 240
Db 181 EWTVLTLYRHKQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFL 240

Qy 241 KEGERHLVSKLYTDMRPETVETTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTSPE 300
Db 241 KEADRLHVSLTYTDMRPETVETTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTSPE 300

Qy 301 DMMWQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMWQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360

Qy 361 TRACRIIPAFHGWNVYKCHACCVLNROGGKRNVLPHGPASTGKSIITAOAIAQAVGNV 420
Db 361 TRTCRIIPAFHGWNVYKCHACCVLNROGGKRNVLPHGPASTGKSIITAOAIAQAVGNV 420

Qy 421 CYNAAVNFNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTIIRIDQKGSQIEPTP 480
Db 421 CYNAAVNFNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTIIRIDQKGSQIEPTP 480

Qy 481 VIMTTNENITVVRIGCEBERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAMLVK 540
Db 481 VIMTTNENITVVRIGCEBERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAMLVK 540

Qy 541 NGYQSTWASACAKWKGVDPDMSENWAEKVPPIINLLGSARSPFTTPKSTPLSQNYALTPL 600
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds  
(without alignments)  
11177.792 Million cell updates/sec

Title: US-10-069-056-12  
Perfect score: 2019  
Sequence: 1 atggctggaatgcttactc.....agccgtgaacttggaactaa 2019

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2019	100.0	2019	6	AX137747 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	V01115 Minute viru
4	2017.4	99.9	5149	14	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137739 Sequence
6	2015.8	99.8	2019	6	AX137743 Sequence
7	2015.8	99.8	2019	6	AX137751 Sequence
8	1905.4	94.4	5085	14	M03322 Minute viru
9	1902.2	94.2	4764	14	M03323 Minute viru
10	1902.2	94.2	5087	14	X02481 Mouse parvo
11	1899	94.1	5144	14	U12469 Mouse parvo
12	1873	92.8	4764	14	M03324 Minute viru
13	1862.2	92.2	4773	14	H0334255 Mouse parvo
14	1855.8	91.9	4761	14	M0334256 Mouse parvo
15	1771	87.7	5121	6	CQ786765 Mice minute
16	1662.2	82.3	4904	14	AF321230 Sequence
17	1659	82.2	4795	14	AF322884 Rat minut
18	1657.4	82.1	4927	14	U79033 Kilham rat
19	1656.6	82.1	5135	14	M81888 Parvovirus

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARH1	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	51.8	5124	14	PVPFVC	M38246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCY1A	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAPFVNS1	X55115 Feline panl
40	1043.6	51.7	5124	14	PVCPVC	M38245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVMRFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137747	AX137747	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	Sequence	12 from Patent	EP1077260.			
DEFINITION	Sequence	12 from Patent	EP1077260.			
ACCESSION	AX137747					
VERSION	AX137747.1	GI:14273921				
KEYWORDS	Mice minute virus					
SOURCE	Mice minute virus					
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.					
REFERENCE	1					
AUTHORS	Nuesch, J. and Rommelaere, J.					
TITLE	Parvovirus ns1 variants					
JOURNAL	Patent: EP 1077260-A 12 21-FEB-2001;					
	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts					
	(DE)					
FEATURES	Location/Qualifiers					
source	1. .2019					
	/organism="Mice minute virus"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:10794"					
	/note="Parvovirus NS1 variant"					
CDS	1. .2019					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAC39995.1"					
	/db_xref="GI:14273921"					
	/translation="MAGNAYSDEVLTGATNWLKEKSNQEVFSFVKENVNOLNGKDIGM					
	NSYKLEQDELKSLQRGAEITWDQSEDMEWETTVDMTKKQVFIQDLSLVKKCLFEVL					
	NTKNIFPGDVNMFVHEWGDQGHCHVLGGKDFSAQKQGWRRQLNVTWSRWLVTA					
	CNVLPAPRIKLREIAEDNFWLLTYHKQTKDYTKCVLFGNMAYVFLTKKIS					
	TSPRRGGYFLSDSGKTNFKLGERHLVSKLYTDMDRPTETVTTVTAQETKRGRI					
	QTKKYSIKITLKLKELVHRTSPEDMMWQPSYIEMMAQGGENLKNLEICTLTFL					
	ARTKAPDLILEKATSKLTNPSLPTDTCRIEAFHGWNVKVCNACCVLNRQGGKR					
	NAVLFKGPASTGKSIITAAIAQAVNGVGCYNAANVPFNDCTNKNLWVEEAGNFQ					
	QNVQKAIKSGTIRIDKQKSKQIEPTVIMTNITVTRIGCEERPEHTOPTIRD					
	RLNLIHLTLFGDFGLVKNWPMICAWLVKNGYQSTWASACAKGKQEDWSENWAE					
	PKVPTINILGARSPTTPKSTPLSQNTALTPLASDLEDALEPSTNTPTVAGTAE					
	TQNTGAGSKACQDGLSPTWSIEEDLRACFGAEPLKXDFSEPLNLD"					

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCTGGAATGCTTACTCTGATGAAGTCTTTGGAGCAACCAACTGGTTAAAGGAAAA	60				
Ds	1	ATGCTGGAATGCTTACTCTGATGAAGTCTTTGGAGCAACCAACTGGTTAAAGGAAAA	60				
Qy	61	AGTAACCAAGGAGTCTTCTCAATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGGAAAA	120				
Ds	61	AGTAACCAAGGAGTCTTCTCAATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGGAAAA	120				
Qy	121	GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGCGAGCTGAAATCTTTTCAA	180				
Ds	121	GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGCGAGCTGAAATCTTTTCAA	180				
Qy	181	CGAGAGCGGAAACTTACTTGGGACCAAGCAGGACATGGAATGGGAAACACACAGTGGAT	240				
Ds	181	CGAGAGCGGAAACTTACTTGGGACCAAGCAGGACATGGAATGGGAAACACACAGTGGAT	240				
Qy	241	GAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA	300				
Ds	241	GAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA	300				
Qy	301	GTGCTTAAACAAAGATATATTTCTGGTGAATGTTTAAATGTTTGTGCAACATGAATGG	360				
Ds	301	GTGCTTAAACAAAGATATATTTCTGGTGAATGTTTAAATGTTTGTGCAACATGAATGG	360				
Qy	361	GGAAAGACCAAGCTGCGACATGCTAATTCGAGGAAAGGACTTTAGTCAAGCT	420				
Ds	361	GGAAAGACCAAGCTGCGACATGCTAATTCGAGGAAAGGACTTTAGTCAAGCT	420				
Qy	421	CAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC	480				
Ds	421	CAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC	480				
Qy	481	TGTAATGTCAACTAAACACAGCTGAAAGATTTAACTAAGAGAAATAGCAGAGCAAT	540				
Ds	481	TGTAATGTCAACTAAACACAGCTGAAAGATTTAACTAAGAGAAATAGCAGAGCAAT	540				
Qy	541	GAGTGGGTACTCTACTTATTAAGCATAAGCAACCAACCAAAAGACTATACCAAGTGT	600				
Ds	541	GAGTGGGTACTCTACTTATTAAGCATAAGCAACCAACCAAAAGACTATACCAAGTGT	600				
Qy	601	GTCTTTTGGAAACATGATGTTTACTATTTTAACTAAAGAAAAATTAAGCCTAGT	660				
Ds	601	GTCTTTTGGAAACATGATGTTTACTATTTTAACTAAAGAAAAATTAAGCCTAGT	660				
Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTA	720				
Ds	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTA	720				
Qy	721	AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGCCAGAAACG	780				
Ds	721	AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGCCAGAAACG	780				
Qy	781	GTTGAAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTTCAAACTAAAAA	840				
Ds	781	GTTGAAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTTCAAACTAAAAA	840				
Qy	841	GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAG	900				
Ds	841	GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAG	900				
Qy	901	GACTGGATGATGTCAGCAGCAGAGTTTACATTTAAATGATGGCTCAACAGGTGGAGAA	960				
Ds	901	GACTGGATGATGTCAGCAGCAGAGTTTACATTTAAATGATGGCTCAACAGGTGGAGAA	960				
Qy	961	AACCTGTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAACCAAAACAGCA	1020				
Ds	961	AACCTGTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAACCAAAACAGCA	1020				
Qy	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCACTAACCACTTTTCACTGCCTGAC	1080				

RESULT 2  
AX137736  
LOCUS  
DEFINITION  
ACCESSION  
VERSION

AX137736  
Sequence 1 from Patent EP1077260.  
AX137736  
AX137736.1 GI:14273909

Ds	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAAACCAACTTTTCACTGCCTGAC	1080
Qy	1081	ACAGAACCTGCGAGATTTTGTCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT	1140
Ds	1081	ACAGAACCTGCGAGATTTTGTCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT	1140
Qy	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTATTCATGGACCA	1200
Ds	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTATTCATGGACCA	1200
Qy	1201	GCAGACAGGCAAAATCTATTATGCAAGCCATAGCAAGAGTGGCAATGTTGGT	1260
Ds	1201	GCAGACAGGCAAAATCTATTATGCAAGCCATAGCAAGAGTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATGCAAGCAATGTAATCTTTTCAATTTTAAAGTGTACCAACAGAACTTGATT	1320
Ds	1261	TGCTATAATGCAAGCAATGTAATCTTTTCAATTTTAAAGTGTACCAACAGAACTTGATT	1320
Qy	1321	TGGGTAGAAGAGCTGGTAACTTTTGACAGCAAGTAAACCAAGTAAAGCCATTTGCTCT	1380
Ds	1321	TGGGTAGAAGAGCTGGTAACTTTTGACAGCAAGTAAACCAAGTAAAGCCATTTGCTCT	1380
Qy	1381	GGTCAAACTATTTCGATTTGATCAAAAGGAAAGGAGCAAGAAAGATTTGAACCAACCA	1440
Ds	1381	GGTCAAACTATTTCGATTTGATCAAAAGGAAAGGAGCAAGAAAGATTTGAACCAACCA	1440
Qy	1441	GTCAATGATGACCAAAATGAGAACATTTACAGTGGTCAAGTGGCTGCGAAGAAAGACCA	1500
Ds	1441	GTCAATGATGACCAAAATGAGAACATTTACAGTGGTCAAGTGGCTGCGAAGAAAGACCA	1500
Qy	1501	GAAACACACTCAACCAATCGAGACAGAAATGCTTTAACTTATCATCAACATACCTTGCCT	1560
Ds	1501	GAAACACACTCAACCAATCGAGACAGAAATGCTTTAACTTATCATCAACATACCTTGCCT	1560
Qy	1561	GGTCACTTGGTTGGTTGACAAAAATGAATGSCCCATGATTTGCTGGTGGTAAAG	1620
Ds	1561	GGTCACTTGGTTGGTTGACAAAAATGAATGSCCCATGATTTGCTGGTGGTAAAG	1620
Qy	1621	AATGGTTTCACTTACCATGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATGG	1680
Ds	1621	AATGGTTTCACTTACCATGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATGG	1680
Qy	1681	TCAGAAAACTGGCGGAGCAAAAGTGGCCAACTCTTATAAATTTACTAGGTTCGGCAGCG	1740
Ds	1681	TCAGAAAACTGGCGGAGCAAAAGTGGCCAACTCTTATAAATTTACTAGGTTCGGCAGCG	1740
Qy	1741	TCACCACTTCAAGCACCGAAAGTACGCTCTCAGCCAGAACTATGCATTAACCTCCTT	1800
Ds	1741	TCACCACTTCAAGCACCGAAAGTACGCTCTCAGCCAGAACTATGCATTAACCTCCTT	1800
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTCTTGGG	1860
Ds	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTCTTGGG	1860
Qy	1861	GGCAGTGCAGAAACCCAGAACCACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA	1920
Ds	1861	GGCAGTGCAGAAACCCAGAACCACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA	1920
Qy	1921	CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTCGGGAACCG	1980
Ds	1921	CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTCGGGAACCG	1980
Qy	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019	
Ds	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019	

[illegible]

```
QY 1621 AATGTTTACCAATCTACCATGCGAAGCTACTGTCTAAATGGGCAAAAGTTCCTGATGG 1680
D 1621 AATGTTTACCAATCTACCATGCGAAGCTACTGTCTAAATGGGCAAAAGTTCCTGATGG 1680
QY 1691 TCAGAAACTGGGCGGAGCCAAAGTGCGCACTCTCTATAAATTTACTAGTTGCGCACGC 1740
D 1691 TCAGAAACTGGGCGGAGCCAAAGTGCGCACTCTCTATAAATTTACTAGTTGCGCACGC 1740
QY 1741 TCACATTTCAGACACCGCAAAAGTACGCTCTCAGCAGAACTATGCACTAACTCCACTT 1800
D 1741 TCACATTTCAGACACCGCAAAAGTACGCTCTCAGCAGAACTATGCACTAACTCCACTT 1800
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCACACCAATATCTCTGTTGCG 1860
D 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCACACCAATATCTCTGTTGCG 1860
QY 1861 GGCCTGTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920
D 1861 GGCCTGTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920
QY 1921 CTGAGCCCACTTCGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
D 1921 CTGAGCCCACTTCGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAATCTTGACTAA 2019
D 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAATCTTGACTAA 2019

RESULT 3
PAMV2
LOCUS PAMV2 5081 bp DNA linear VRL 10-FEB-1999
DEFINITION Minute virus of mice with two major open reading frames (genome).
ACCESSION V01115
VERSION V01115.1 GI:60911
KEYWORDS coat protein; genome; origin of replication; overlapping genes;
terminal repeat.
SOURCE
ORGANISM
Mice minute virus
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 5081)
Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
The complete DNA sequence of minute virus of mice, an autonomous
parvovirus
Nucleic Acids Res. 11 (4), 999-1018 (1983)
JOURNAL
MEDLINE 83143341
PUBMED 6298737
COMMENT The messenger RNA of this virus (colinear with the strand listed)
is spliced, but the exact splice sites are not known. The listed
strand is complementary to the one which is included in the virion.
FEATURES
Location/Qualifiers
1..5081
/organism="Mice minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
114..2279
/feature="unlabeled protein product; coding sequence"
/codon_start=1
/protein_id="CAA24309.1"
/db_xref="GI:60912"
/db_xref="GOA:P01134"
/db_xref="Swiss-Prot:P01134"
/translation="MISGSGSLNQKGRKWAWEKVKYKOLLKSVYLLFFHSVSRDAOKE
SNQLTMAGNAYSEVIGATNLKESNQEVFSFVKFNENVQLNGKIDGNYSYKEIQE
DELKSLQRAETWDQSEMETTVDKMTKQVIFDLSLVKCLFVLNKNIFPGD
VNFVFEWQKQDQCHVLIGKDFSAQGRWRRQLNLYWRSVLVATACNQLTAE
RIKLRETAEDNEWTLTYKHQTKDYTKCVLFGNMAYFLTKKIKSTSPRDGGY
LSSDSGWKTNFLKGRHLVSKLYDDMRPVTETVTTAQTKEGRIQTKVEISIK
TLKELVHKVTSPEDDMMQPSYLEMMAQPGENLLKNTLIEITLARTKTAFDL
ILLEAETSKLNTSLPRTCRIFAFHGWYKVKCHACVLRNQGKNTVLPHQPA
SGKSIITQAIAQAGNVGCYNANVFPNDCTNKNLIWVEAGNFGQGVNQFKIHC
SGOTIRIDQKGRGSKQIEPTFVIMTNNENITVVRIGCEBERPEHTQPIRDMNLHLTH
```

## ORIGIN

```
Query Match 99.9%; Score 2017.4; DB 14; Length 5081;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAGGAAAAA 60
D 261 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAGGAAAAA 320
QY 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGCTTCAACTGAATGAAAA 120
D 321 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGCTTCAACTGAATGAAAA 380
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180
D 381 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 440
QY 181 CGAGGAGCGGAAATCTCTTGGGACCAACGAGGACATGGAATGGGAAACCAAGTGGAT 240
D 441 CGAGGAGCGGAAATCTCTTGGGACCAACGAGGACATGGAATGGGAAACCAAGTGGAT 500
QY 241 GAAATGACCAAAAAAGCAAGTATTTCTTTTGTGTTTAAAAAATGTTTATTTGAA 300
D 501 GAAATGACCAAAAAAGCAAGTATTTCTTTTGTGTTTAAAAAATGTTTATTTGAA 560
QY 301 GTGCTTAAACACAAAGAATAATTTCTCTGTGATGTTAATTTGTTGTGCAACATGAATGG 360
```

```
TLPDFGLVDKNEWPMICAWLVKNGYQSTWASYCAKWKGVDPWSENWABEKVPPTINL
LGSARPFPTPKSTPLSQNYALTPLASDLIEDLALPEWSTENTPVAGTAETQNTGAGS
KACQDQSLPTWSEIEEDLRACFGAEPLKDFSEPLND"
2286..2254
/notes="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24310.1"
/db_xref="GI:60913"
/db_xref="GOA:P01137"
/db_xref="Swiss-Prot:P01137"
/translation="MAPPAKRAKRGKLRDGLWLVGY"
2332..2361
/notes="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24311.1"
/db_xref="GI:584078"
/db_xref="TrEMBL:Q89491"
/translation="MVGWNGIN"
2354..2398
/notes="unnamed protein product; coding sequence"
/codon_start=1
/protein_id="CAA24312.1"
/db_xref="GI:60914"
/db_xref="TrEMBL:Q89818"
/translation="MENLYFYRPEITWF"
2383..4551
/notes="unnamed protein product; Protein sequence is in
conflict with the conceptual translation; coding sequence"
/codon_start=1
/protein_id="CAA24313.1"
/db_xref="GI:1335397"
/db_xref="GOA:Q84207"
/db_xref="TrEMBL:Q84207"
/translation="NHLVLGWPPGYKYLPGNSLDQGEPTNPSDAAAKEHDEAYDQY
IKSKNPYLIFSAADQRFIDQTKADKMGKGVHYFFRTKRAFAPKLATDSBEGTSGV
SRAGKTRPPAYIFINQARAKKLTSSAAQSQSDTSDGSDPSDNVHSAARVERA
ADPGGGGGGGGGVSTGSDNQTHYRFLGDCWEITATLRLVHLNPKSERY
CIRVHNTTDSVKGNKXDAHQEIWTWPSLVADANAGVWLQSPDWQVICTMSOLN
LVSLDOEIIENVLKTVEODLGOAIIKYNNDLTACMVAVDSNNILPVTPAANSMET
LGFYWPKPIASPYRYFCVDRDLSVTYENQSGTVEHNVMTGPKGIPQFTIENTQOI
TLRTGDEFATGTTFDTSVKLTHTFCQPNDFEASRAGPFAAPKVPADITAGDTLTAQSGR
HGTQMGVNVSEAIRPAQVGFQPHNDFEASRAGPFAAPKVPADITAGDTLTAQSGR
SVRYSGKHGHNWASHGAPRYTWDSTFSGRDQKDFIQSAPLVVPPPLINGILT
NANPICKNDIHFSNVFNSGPLTADSPHSVPVPOQIWDKELDLHKRLHITADPV
CKNAPGQMLVRLGNLTDQYDPNGATLSRIVTGYTFWKKGLTWAKLRANTTNWNPV
YQVSAEDNGNSYMSVTKLPTATGNMQSVPLITRPVARNY"
```

|||||  
561 GTGCTTAACAAAGAAATATTTCTCTGGTGTATGTTAAATGGTTTGTGCAACATGAATGG 620  
|||||  
361 GGAAGAGACCAAGCGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTTACTCAAGCT 420  
|||||  
621 GGAAGAGACCAAGCGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTTACTCAAGCT 680  
|||||  
421 CAAGGGAATTTGGTGGAGAGGCAATTAATTTTACTTGGAGCAGATGGTTGGTAAAGCC 480  
|||||  
681 CAAGGGAATTTGGTGGAGAGGCAATTAATTTTACTTGGAGCAGATGGTTGGTAAAGCC 740  
|||||  
481 TGTAAATGTCAACTAATACACAGCTGAAGAAATTAACAAAGCAATAGCAGAAAGCAAT 540  
|||||  
741 TGTAAATGTCAACTAATACACAGCTGAAGAAATTAACAAAGCAATAGCAGAAAGCAAT 800  
|||||  
541 GAGTGGGTACTTACTTACTTATAGCATAAGCAACCAAAAAAGCAATATACCAAGTGT 600  
|||||  
801 GAGTGGGTACTTACTTACTTATAGCATAAGCAACCAAAAAAGCAATATACCAAGTGT 860  
|||||  
601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGCAATATAGCACTAGT 660  
|||||  
861 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGCAATATAGCACTAGT 920  
|||||  
661 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAACTTAACTTTTA 720  
|||||  
921 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAACTTAACTTTTA 980  
|||||  
721 AAAGAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780  
|||||  
981 AAAGAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 1040  
|||||  
781 GTTGAACCAACAGTAACCTAGCGAGGAACTAAGCGCGGAGAAATTCAAAATAAAAA 840  
|||||  
1041 GTTGAACCAACAGTAACCTAGCGAGGAACTAAGCGCGGAGAAATTCAAAATAAAAA 1100  
|||||  
841 GAAGTTTCTAATTAACACTACACTTAAGAGCTGGTGCATTAAGAGTAACCTCACCAG 900  
|||||  
1101 GAAGTTTCTAATTAACACTACACTTAAGAGCTGGTGCATTAAGAGTAACCTCACCAG 1160  
|||||  
901 GACTGGATGATGACAGCAGACAGTTACATTGAATGATGGCTCAACAGGCTGGAGAA 960  
|||||  
1161 GACTGGATGATGACAGCAGACAGTTACATTGAATGATGGCTCAACAGGCTGGAGAA 1220  
|||||  
961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAAACCAACAGCA 1020  
|||||  
1221 AACTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAAACCAACAGCA 1280  
|||||  
1021 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080  
|||||  
1281 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1340  
|||||  
1081 ACAGAACCTGCAGAAATTTTGTCTTTTCTGCTGGAATCTATGTTAAAGTTTGCATGCT 1140  
|||||  
1341 ACAGAACCTGCAGAAATTTTGTCTTTTCTGCTGGAATCTATGTTAAAGTTTGCATGCT 1400  
|||||  
1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1200  
|||||  
1401 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1460  
|||||  
1201 GCCAGCAGAGCAAACTATATTATGCAAGGCCATAGCAAGCAGTTGGCAATTTGGT 1260  
|||||  
1461 GCCAGCAGAGCAAACTATATTATGCAAGGCCATAGCAAGCAGTTGGCAATTTGGT 1520  
|||||  
1261 TGCTATTAATGCGCAATGTAACTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
|||||  
1521 TGCTATTAATGCGCAATGTAACTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1580  
|||||  
1321 TGGGTAGAGAGCTGGTAACTTTGGAGCAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380  
|||||  
1581 TGGGTAGAGAGCTGGTAACTTTGGAGCAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1640  
|||||  
1381 GGTCAAACTATTCGCTATTTGATCAAAAAGGAAAGGCGAGCAAAAGATTTGAACCAACCA 1440  
|||||

Db 1641 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGCAAGATTAACCAACCA 1700  
Qy 1441 GTCATCATGACCAAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTTCGGAAGAAACCA 1500  
Db 1701 GTCATCATGACCAAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTTCGGAAGAAACCA 1760  
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTCAACATACCTTTGCCT 1560  
Db 1761 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTCAACATACCTTTGCCT 1820  
Qy 1561 GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTGTTGCTTTGGTTGTTAAAG 1620  
Db 1821 GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTGTTGCTTTGGTTGTTAAAG 1880  
Qy 1621 AATGTTTCAATATACCATGAGCAAGTACTGTCTTAAATGGGGCAAGTTCTCTGATTGG 1680  
Db 1881 AATGTTTCAATATACCATGAGCAAGTACTGTCTTAAATGGGGCAAGTTCTCTGATTGG 1940  
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAATCTCTATATAATTTACTAGCTTCGGCAGCG 1740  
Db 1941 TCAGAAAACTGGCGGAGCCAAAGGTGCCAATCTCTATATAATTTACTAGCTTCGGCAGCG 2000  
Qy 1741 TCACCAATTCACGACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCCACTT 1800  
Db 2001 TCACCAATTCACGACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCCACTT 2060  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCAGCAACCAATACTCTCTGTTGG 1860  
Db 2061 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCAGCAACCAATACTCTCTGTTGG 2120  
Qy 1861 GGCATCTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAC 1920  
Db 2121 GGCATCTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAC 2180  
Qy 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 1980  
Db 2181 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 2240  
Qy 1981 TTGAAGAAAGACTTTCAGCAGCGCTGAACTTGGACTAA 2019  
Db 2241 TTGAAGAAAGACTTTCAGCAGCGCTGAACTTGGACTAA 2279

## RESULT 4

WMVPCG 5149 bp ss-DNA linear VRL 22-MAY-1995  
LOCUS Minute virus of mice, complete genome.  
DEFINITION J02275 M12520 M12521 M14704  
ACCESSION J02275.1 Gi:332293  
VERSION alternative splicing; capsid protein; complete genome;  
KEYWORDS nonstructural protein.  
SOURCE Mice minute virus  
ORGANISM Mice minute virus  
REFERENCE 1 (bases 1 to 5149)  
AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.  
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus  
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)  
MEDLINE 83143341  
PUBMED 6298737  
REFERENCE 2 (bases 1 to 5149)  
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.  
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain  
J. Virol. 57 (2), 656-669 (1986)  
MEDLINE 86115415  
PUBMED 3502703  
REFERENCE 3 (sites)  
AUTHORS Morgan,W.R. and Ward,D.C.  
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs



CDS	2354..2398 /gene="vp" /note="ORF2; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67113.1" /db_xref="GI:825484" _translation="MFNYLFYRPEITWF" 2399..>4557 /gene="vpi" /number=2  2794..4557 /gene="vpi" /note="vp2" /codon_start=1 /protein_id="AAA67114.1" _translation="MSDCTSQPSGNVHSAARVERAADPGSGGGGCGGCVGVST GSYDQIQTFLPGLDGWVEITATLALRLVHLNMPKSENCRIRVHTTIDTSVKGNMAKDD AHEQIQTPWSLNDLANGWLPQSDMQYICNTISQLNLSDQEIENVLTKTYEODL GGQAIKLYNLDLTACMMVAVDSSNNILFYTPAANSMETLGFYPWKPTIASPYRYFCVD RDLSVTVENEGTEVHNVMGTPKMNSQFFTIENTQOITLLRTGDFATGYTYVFDTNS VKLTHTWQTRQGLQPPLLSTFEADTDAGTLTAQGRHGTTOMGVNWVSEAIRTERPA QVGCFPHNDFEASRGTFAPAKVPADITQGVDEKANGSVRSYIGKHQENWASHGPA PERYTWDTSFGGRDKPFQIOSAPLVPPPLNGILTANPIGTKNDIHFSNVFNYSY GPLTAFSHSPSPVPOQQIWDBKELDLHKPRLIHTAPFVCXKNAPQGMLVRLGNLTDQ
exon	
CDS	
Query Match	99.9%; Score 2017.4; DB 14; Length 5149;
Best Local Similarity	100.0%; Pred. No. 0;
Marches 2018; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 ATGCCTGGAAATGCTTACTCTGATGAAGTTTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Dd	261 ATGCTGGAAATGCTTACTCTGATGAAGTTTTTGGGAGCAACCACTGGTTAAAGGAAAA 320
Qy	61 AGTAAACAGGAAGTGTTCTCATTTGTTTTTAAAAATGAAAATGTTCAACTGAATGGAAAA 120
Dd	321 AGTAAACAGGAAGTGTTCTCATTTGTTTTTAAAAATGAAAATGTTCAACTGAATGGAAAA 380
Qy	121 GATAATCGATGGAATAGTTACAAAAAGAGCTGCAGGAGCAGCAGCTGAAATCTTTACAA 180
Dd	381 GATAATCGATGGAATAGTTACAAAAAGAGCTGCAGGAGCAGCAGCTGAAATCTTTACAA 440
Qy	181 CGAGGAGCGGAAACTACTTGGGACCAAAGCGAGGACATGGAATTGGGAAACCACAGTGGAT 240
Dd	441 CGAGGAGCGGAAACTACTTGGGACCAAAGCGAGGACATGGAATTGGGAAACCACAGTGGAT 500
Qy	241 GAAATGACCAAAAAGCAAGTATTCAATTTTGATTCCTTTGGTTAAAAATGTTTATTGGAA 300
Dd	501 GAAATGACCAAAAAGCAAGTATTCAATTTTGATTCCTTTGGTTAAAAATGTTTATTGGAA 560
Qy	301 GTGCTTTAACACAAGAATATATTTCCCTGGTGATGCTTAATTTGGTTGTGCAACATGAATGG 360
Dd	561 GTGCTTTAACACAAGAATATATTTCCCTGGTGATGCTTAATTTGGTTGTGCAACATGAATGG 620
Qy	361 GGAAAAGACCAAGGCTGGCACTGCCATGTACTAAATTGGAGAAAGGACTTTTAGTCAAGCT 420
Dd	621 GGAAAAGACCAAGGCTGGCACTGCCATGTACTAAATTGGAGAAAGGACTTTTAGTCAAGCT 680
Qy	421 CAAGGGAATGGTGGGAAGGCAACTAAATGTTTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Dd	681 CAAGGGAATGGTGGGAAGGCAACTAAATGTTTTTACTGGAGCAGATGGTTGGTAAACAGCC 740
Qy	481 TGTAAATGTGCNACTAAACACCAAGCTGAAGAATTAATACTAAGAGNAATAGCAGAGACAAT 540
Dd	741 TGTAAATGTGCNACTAAACACCAAGCTGAAGAATTAATACTAAGAGNAATAGCAGAGACAAT 800
Qy	541 GAGTGGGTTACTCTTACTTACTTATAAGCATGAAGCAAAACCAAAAAGACTATACCAAGTGT 600
Dd	801 GAGTGGGTTACTCTTACTTACTTATAAGCATGAAGCAAAACCAAAAAGACTATACCAAGTGT 860
Qy	601 GTTCTTTTGGAAACATGATTGCTTACTATTATTTTTTAACTAAAAAGAAAAAAGCACTAGT 660



```
QY 1741 TCACATTTCAGACACCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
DB 2001 TCACATTTCAGACACCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAAACTATCTCTGTTGCG 1860
DB 2061 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAAACTATCTCTGTTGCG 2120
QY 1861 GGCACCTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA 1920
DB 2121 GGCACCTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA 2180
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTCTCGGTGGGAAACCG 1980
DB 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTCTCGGTGGGAAACCG 2240
QY 1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGACTAA 2019
DB 2241 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGACTAA 2279

RESULT 5
AX137739 LOCUS AX137739 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 4 from Patent EP1077260.
ACCESSION AX137739
VERSION AX137739.1 GI:14273913
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1
AUTHORS Nuesch, J. and Rommelaere, J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 4 21-FEB-2001;
DEUTSCHES Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
Source location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1..2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273914"
/translation="MAGNAYSDEVLGATNMLKEKSNOEVFVFKENVQLNGKDIGW
NSYKKELOEDELKSLQSGAETTWQSEDMEWETVDEMTKKQVFI PDSL VKKCLPEVL
NTKNI PFGDVNWFVHEWKGDKGCHVLIGKDFSOAGKWRRLNLYVNSRLVTA
CNVQLTPAERI KLEIAEDNEWVLLTYKHQTKDYTKCVLFGNMIAFYFLTKKIS
TSPRDGGYFLSSDSGKNTPLKEGKHLVSKLYTDDMRPEVETITVTTAQETPKRI
QTKKEVAIKITLKHVRVTSPEDEMMQPDYSIEMMAQPGGENLLKNTLEICTLTL
ARTKPAFLILEAKETSKLNFSLPDTTRCRI FAFHGNVYKVCCHAI ICVLRNQGGR
NTVLFHPGASTGKSI IAOAQAQVGNVCYNAANVPFPNDCTNKL I WVEBAQNGFQ
QVNFKAICSGQOTRIDQKGSKOIEPTVPVINTNENITVVRIGCEERPEHTQPIRD
RMLNIHLTHLPGDFGLVDKNEPMICAMLVKNQYQSTMASYCAKWKVQPDWSENWAE
PKYPTPINLGSARSPTTFSKTSQNYALTPLASDLEDLADEPWSSTENTPVAGTAE
TQNTGEAGSKACQDQGLSPSTWSEIEEDLRACFGAEPLKDFSEPLNLD"

ORIGIN
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCTCGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60
DB 1 ATGCTCGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60
QY 61 AGTAACGAGAGGTTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
```

```
DB 61 AGTAACGAGAGGTTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
QY 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGAGRGAGCAGCTGGAATCTTTTACAA 180
DB 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGAGRGAGCAGCTGGAATCTTTTACAA 180
QY 181 CGAGGAGCGGAAATACTCTTTGGGACCAAGCAGGACATGGAATGGGAAACCAACAGTGGAT 240
DB 181 CGAGGAGCGGAAATACTCTTTGGGACCAAGCAGGACATGGAATGGGAAACCAACAGTGGAT 240
QY 241 GAAATGACCAAAAAAGCAAGATTTCAATTTTGTATCTTTTGTGTTAAAAAATGTTTATTGAA 300
DB 241 GAAATGACCAAAAAAGCAAGATTTCAATTTTGTATCTTTTGTGTTAAAAAATGTTTATTGAA 300
QY 301 GTCTTTAAACAAGAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCAACATGAATGG 360
DB 301 GTCTTTAAACAAGAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTTAGTCAAGCT 420
DB 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTTAGTCAAGCT 420
QY 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGAGCAGATGGTTGGTAAACAGCC 480
DB 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGAGCAGATGGTTGGTAAACAGCC 480
QY 481 TGTAAATGTGCAACTAAACACGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGCAAT 540
DB 481 TGTAAATGTGCAACTAAACACGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGCAAT 540
QY 541 GAGTGGGTTTACTCTACTTATATAGCATAAGCAACCAACCAAGACTATATACCAAGTGT 600
DB 541 GAGTGGGTTTACTCTACTTATATAGCATAAGCAACCAACCAAGACTATATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATGTTGTTACTATTTTTTAACTAAAGAAATATAGCACTAGT 660
DB 601 GTTCTTTTGGAAACATGATGTTGTTACTATTTTTTAACTAAAGAAATATAGCACTAGT 660
QY 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTGACTCTGGCTGGGAAACTAACTTTTAA 720
DB 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTGACTCTGGCTGGGAAACTAACTTTTAA 720
QY 721 AAAGAGGCGAGCGCCCATCTAGTGAGCAAACTATACACTGATGACATCGGCGCAGAAACG 780
DB 721 AAAGAGGCGAGCGCCCATCTAGTGAGCAAACTATACACTGATGACATCGGCGCAGAAACG 780
QY 781 GTTGAACACACAGTAACCACTGGCGAGGAAACTAAGCGCGGAGAAATTCAAACTAAAAA 840
DB 781 GTTGAACACACAGTAACCACTGGCGAGGAAACTAAGCGCGGAGAAATTCAAACTAAAAA 840
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGTCATAAAAAGAGTAACCTCACCAGAG 900
DB 841 GAAGTTGCTATTAAAACTACCTTAAAGAGCTGGTGTCATAAAAAGAGTAACCTCACCAGAG 900
QY 901 GACTGGATGATGATGAGCAGCAGACAGTTTACATTGAAATGATGGCTCAACCGAGTGGAGAA 960
DB 901 GACTGGATGATGATGAGCAGCAGACAGTTTACATTGAAATGATGGCTCAACCGAGTGGAGAA 960
QY 961 AACCTGCTGAAAAATACCGTAGAGATTTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
DB 961 AACCTGCTGAAAAATACCGTAGAGATTTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
QY 1021 TTTGACTTAATTTTAAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC 1080
DB 1021 TTTGACTTAATTTTAAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC 1080
QY 1081 ACAAGAACCTGAGAAATTTTCTTTTATGCTGGAATATGTTTAAAGTTTGCATGCT 1140
DB 1081 ACAAGAACCTGAGAAATTTTCTTTTATGCTGGAATATGTTTAAAGTTTGCATGCT 1140
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTTCATGGACCA 1200
DB 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTTCATGGACCA 1200
```



Qy	1201	GCAGCAGGCAAACTCTATTATGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCAGCAGGCAAACTCTATTATGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATCAGCCATGTAATCTTCCATTTAATGACGTGTACCAACAGACTTGATT	1320
Db	1261	TGCTATAATCAGCCATGTAATCTTCCATTTAATGACGTGTACCAACAGACTTGATT	1320
Qy	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGATTTAAAGCCATTGCTCT	1380
Db	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGATTTAAAGCCATTGCTCT	1380
Qy	1381	GGTCAAACTATTCGATTTGATCAAAAAGGAAAGGAGCAAGATTTGAACCAACCA	1440
Db	1381	GGTCAAACTATTCGATTTGATCAAAAAGGAAAGGAGCAAGATTTGAACCAACCA	1440
Qy	1441	GTCAATCATGACCAAAATGAGACATTTACAGTGTGTAGATAGGCTGCGAAGAACCA	1500
Db	1441	GTCAATCATGACCAAAATGAGACATTTACAGTGTGTAGATAGGCTGCGAAGAACCA	1500
Qy	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCTAACACATACCTTGCCT	1560
Db	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCTAACACATACCTTGCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTCACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTCACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG	1620
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAGTTCTCTGATTGG	1680
Db	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAGTTCTCTGATTGG	1680
Qy	1681	TCAGAAAACCTGGCGGAGCCAAAGGTCCCAACTCTTATAAATTTACTAGGTTGGCAGCG	1740
Db	1681	TCAGAAAACCTGGCGGAGCCAAAGGTCCCAACTCTTATAAATTTACTAGGTTGGCAGCG	1740
Qy	1741	TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCATTAATCCCACTT	1800
Db	1741	TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCATTAATCCCACTT	1800
Qy	1801	GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTTGG	1860
Db	1801	GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTTGG	1860
Qy	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Db	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Qy	1921	CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCCGTTGCGGAACCG	1980
Db	1921	CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCCGTTGCGGAACCG	1980
Qy	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA	2019
Db	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA	2019

RESULT 6  
AX137743  
LOCUS  
DEFINITION  
Sequence 8 from Patent EP1077260.  
ACCESSION  
AX137743  
VERSION  
AX137743.1 GI:14273917  
KEYWORDS  
Mice minute virus  
SOURCE  
ORGANISM  
Mice minute virus  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
1  
REFERENCE  
Nuesch, J. and Rommelaere, J.  
AUTHORS  
Parvovirus ns1 variants  
TITLE  
Patent: EP 1077260-A 8 21-FEB-2001;  
JOURNAL  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

(DB)

FEATURES	Location/Qualifiers
source	1. .2019
	/organism="Mice minute virus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:10794"
	/notes="Parvovirus NS1 variant"
CDS	1. .2019
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC39993.1"
	/db_xref="GI:14273918"
	/translation="MAGNAYSDEVLTGATNWLKEKSQEVESFVKENVQLNGKDIGH
	NSYKLEQLDELKSLQORGAETTDQSDMEWETTVDEMTEKQVIFDLSLVKCLFEVL
	NKNTIPPGVNVNFVQHEWGDQGHVLIIGKDFSOAQKWRRLNVYWSLWYTA
	CNVQTPADRIKLREIAEDNEWVLLTYKHQTKDYTKCVLFNNIAYYFLPKKIS
	TSPPRGVFLRSIDGSKTNFLKEGRLVSKLYTDMRPETVETVTTVAQETKRGRI
	QTKKESVINKTLKELVHKVTSPEMMMPQDSYIEMMAQPGGNNLLKNTLICTLTL
	ARTKAPDLILLEKETSKLNTSLPDRACRIPAFHGNVYKVCCHALCCVLRQGGKR
	NVLPHGPASTGKSIIAQAIQAVGNVGCNNAVNFNFNDCTNKNLILWEEAGNFGQ
	QVNFKAICSGQIRIDOKGSKQIEPTFVINTNENITVVRIGCEERPEHTQPIRD
	RMLNIHLTHLPDGLVDKNWPMICAWLVKNYQSTWASVCAKMKVDPDSENNAE
	PKVPTPINLLGSARSPTTPKSTPLSONVALTPLASDLDELALPEWSTPNTPVAGTAE
	TQNTGEAGSKACODGQLSPTWSEIEEDLRACFAGELPKKDFSEPLND"

ORIGIN

Query Match	99.8%;	Score	2015.8;	DB	6;	Length	2019;
Best Local Similarity	99.9%;	Pred. No.	0;				
Matches	2017;	Conservative	0;	Mismatches	2;	Indels	0;
				Gaps	0;		
Qy	1	ATGGCTGGAAATGCTTACTCTGATGAAGTCTTGGGAGCAACCAACTGGTTTAAAGGAAAA	60				
Db	1	ATGGCTGGAAATGCTTACTCTGATGAAGTCTTGGGAGCAACCAACTGGTTTAAAGGAAAA	60				
Qy	61	AGTAACACGAGAAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGGAAAA	120				
Db	61	AGTAACACGAGAAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGGAAAA	120				
Qy	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180				
Db	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180				
Qy	181	CGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAAACCAAGTGGAT	240				
Db	181	CGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAAACCAAGTGGAT	240				
Qy	241	GAAATGACCAAAAGCAAGTATTTCTTTTGATTTCTTGGTTAAAAAATGTTTATTGTAA	300				
Db	241	GAAATGACCAAAAGCAAGTATTTCTTTTGATTTCTTGGTTAAAAAATGTTTATTGTAA	300				
Qy	301	GTGCTTTAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTCAACATGAATGG	360				
Db	301	GTGCTTTAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTCAACATGAATGG	360				
Qy	361	GGAAAAAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAGTCAAGCT	420				
Db	361	GGAAAAAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAGTCAAGCT	420				
Qy	421	CAAGGGAATGTTGGAGAGGCAACTTAAATGTTTACTGGAGCAGATGTTTGGTTAAACAGCC	480				
Db	421	CAAGGGAATGTTGGAGAGGCAACTTAAATGTTTACTGGAGCAGATGTTTGGTTAAACAGCC	480				
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGACAAT	540				
Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGACAAT	540				
Qy	541	GAGTGGTTTACTTACTTATTAAGCATAAGCAAAACCAAAAGACTATATACCAAGTGT	600				
Db	541	GAGTGGTTTACTTACTTATTAAGCATAAGCAAAACCAAAAGACTATATACCAAGTGT	600				
Qy	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660				
Db	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660				

QY	661	CCACCAAGAGACGAGGCGCTATTTTCTTAGCAGTGA	CTCTGGCTGGAAAACTAACTTTTTTA	720
DB	661	CCACCAAGAGACGAGGCGCTATTTTCTTAGCAGTGA	CTCTGGCTGGAAAACTAACTTTTTTA	720
QY	721	AAAGAAGCGAGCGGCATCTAGTGAGCAAACTATA	CTACTGATGATCATGCGGCCAGAAACG	780
DB	721	AAAGAAGCGAGCGGCATCTAGTGAGCAAACTATA	CTACTGATGATCATGCGGCCAGAAACG	780
QY	781	GTTGAAACCACAGTAACCACTGCGCAGGAAACTA	AGCGCGCAGAAATCAAACCTAAAAA	840
DB	781	GTTGAAACCACAGTAACCACTGCGCAGGAAACTA	AGCGCGCAGAAATCAAACCTAAAAA	840
QY	841	GAAGTTTCTATTAAAACTACACTTAAAGAGCTG	TGTCATAAAAGAGTAACCTCACAGAG	900
DB	841	GAAGTTTCTATTAAAACTACACTTAAAGAGCTG	TGTCATAAAAGAGTAACCTCACAGAG	900
QY	901	GACTGGATGATGATGCAGCCAGACAGTTTACATT	GAAATGATGGCTCAACACAGGTGGAA	960
DB	901	GACTGGATGATGATGCAGCCAGACAGTTTACATT	GAAATGATGGCTCAACACAGGTGGAA	960
QY	961	AACCTGCTGAAAAATA CGCTAGAGATTTGTAC	TAACTCTAGCCAGAACCAAAACAGCA	1020
DB	961	AACCTGCTGAAAAATA CGCTAGAGATTTGTAC	TAACTCTAGCCAGAACCAAAACAGCA	1020
QY	1021	TTTGACTTAATTTTAGAAAAAGCTGAACACGAAA	CTAACCAACTTTTCACTGCCTGAC	1080
DB	1021	TTTGACTTAATTTTAGAAAAAGCTGAACACGAAA	CTAACCAACTTTTCACTGCCTGAC	1080
QY	1081	ACAAGAACTCGAGAAATTTTTCCTTTTCATGCG	TGGAACATATGTTAAAGTTTGCCATGCT	1140
DB	1081	ACAAGAACTCGAGAAATTTTTCCTTTTCATGCG	TGGAACATATGTTAAAGTTTGCCATGCT	1140
QY	1141	ATTTCCTGTGTTTTAAACAGACAAGGAGGCAAAA	AGAAATGCTGTTTATTTCATGGACCA	1200
DB	1141	ATTTCCTGTGTTTTAAACAGACAAGGAGGCAAAA	AGAAATGCTGTTTATTTCATGGACCA	1200
QY	1201	GCCAGCACAGGCAATCTTAATTTTGCACAGCCAT	AGCAACAGCAGTTGGCAATGTGGT	1260
DB	1201	GCCAGCACAGGCAATCTTAATTTTGCACAGCCAT	AGCAACAGCAGTTGGCAATGTGGT	1260
QY	1261	TGCTATAATGACGCAATGTAAAATTTTCCATTT	TAATGACTGACCAACAAGAACTTGATT	1320
DB	1261	TGCTATAATGACGCAATGTAAAATTTTCCATTT	TAATGACTGACCAACAAGAACTTGATT	1320
QY	1321	TGGGTAGAAGAACTGTGTAACTTTTGGACAGCA	AGTAACCAAGTTTAAAGCCATTTGCTCT	1380
DB	1321	TGGGTAGAAGAACTGTGTAACTTTTGGACAGCA	AGTAACCAAGTTTAAAGCCATTTGCTCT	1380
QY	1381	GGTCAAACTATTCGCATTTGATCAAAAAGGAA	AGGACGACAAACAGATTGAACCAACCA	1440
DB	1381	GGTCAAACTATTCGCATTTGATCAAAAAGGAA	AGGACGACAAACAGATTGAACCAACCA	1440
QY	1441	GTCATCATGACCACAAAATGAGAACATTTACAG	TGGTCAAGATAGGCTCGGAAGAAAGACCA	1500
DB	1441	GTCATCATGACCACAAAATGAGAACATTTACAG	TGGTCAAGATAGGCTCGGAAGAAAGACCA	1500
QY	1501	GAAACACTCAACCAATCAGACAGAAATGCTTAA	CAATTCATCTAACACATACCTTGCCT	1560
DB	1501	GAAACACTCAACCAATCAGACAGAAATGCTTAA	CAATTCATCTAACACATACCTTGCCT	1560
QY	1561	GGTGACTTTGGTTTGTGACAAAATGAAATGGCC	CAATGATTTGTGCTTGTTGGTTAAG	1620
DB	1561	GGTGACTTTGGTTTGTGACAAAATGAAATGGCC	CAATGATTTGTGCTTGTTGGTTAAG	1620
QY	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGT	GCTAAATGGGCAAGTTTCTGTATGG	1680
DB	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGT	GCTAAATGGGCAAGTTTCTGTATGG	1680
QY	1681	TCAGAAAACTGGGCGGAGCCAAAGGTGCCAACT	CTCTATAAATTTTATAGGTTTCGGACGC	1740
DB	1681	TCAGAAAACTGGGCGGAGCCAAAGGTGCCAACT	CTCTATAAATTTTATAGGTTTCGGACGC	1740

Qy	1741	TCACGATTCACGACACCCAGAAAGTACGCCCTCTAGCCGACGAGAACTATGCACTAACTCACA	1801			
Db	1741	TCACGATTCACGACACCCAGAAAGTACGCCCTCTAGCCGACGAGAACTATGCACTAACTCACA	1801			
Qy	1801	GCATCGGATCTCGAGGACCTTGGCTTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTCG	1860			
Db	1801	GCATCGGATCTCGAGGACCTTGGCTTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTCG	1860			
Qy	1861	GGCACTCGAGAAACCCAGAACACACTGGGGAAAGCTGGTTTCCAAAGCCCTGCCAAGATGTC	1920			
Db	1861	GGCACTCGAGAAACCCAGAACACACTGGGGAAAGCTGGTTTCCAAAGCCCTGCCAAGATGTC	1920			
Qy	1921	CTGAGCCCACTTGGTTCAGAGATCGAGGAGATTCGAGAGCTGCTTCGTCGGGAACCG	1980			
Db	1921	CTGAGCCCACTTGGTTCAGAGATCGAGGAGATTCGAGAGCTGCTTCGTCGGGAACCG	1980			
Qy	1981	TTGAAGAAAGACTTCACGAGCGCCCTGAACTTGGACTAA	2019			
Db	1981	TTGAAGAAAGACTTCACGAGCGCCCTGAACTTGGACTAA	2019			
RESULT 7						
AX137751						
LOCUS						
DEFINITION	AX137751	2019 bp	DNA			
ACCESSION	AX137751	Sequence 16 from Patent	Ep1077260.			
VERSION	AX137751.1	GI:14273925				
KEYWORDS	Mice minute virus					
SOURCE	Mice minute virus					
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.					
REFERENCE	1					
AUTHORS	Nuesch, J. and Rommelaere, J.					
TITLE	Parvovirus ns1 variants					
JOURNAL	Patent: Ep 1077260-A 16 21-FEB-2001;					
	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts					
	(DB)					
FEATURES						
source						
1. .2019						
/organism="Mice minute virus"						
/mol_type="unassigned DNA"						
/db_xref="taxon:10794"						
/note="Parvovirus NS1 variant"						
1. .2019						
/note="unnamed protein product"						
/codon_start=1						
/protein_id="CAC39997.1"						
/db_xref="GI:14273926"						
CDS						
1. .2019						
/translation="MAGNAYSDVLTGATNWLKESNQEVFVFKNNVQLNGKDIGW						
NSYKBLQDELKSLQRGAEITWDQSEDMEWETTVDEMTKQVIFPDSLAVKKCLFEVL						
NTXNIIPFGDVFNVFVHGWKQDQGWCHVLIGGKDFSQAGKWRRLQNVYWSRWLVTA						
CNVLQTPAEIRIKLRETAEDNWTLLTYKHQTKKDYTKVLFGNNIAYFUTKKKIS						
TSPRGDGYFLSSDCGWKTNELKEGERHLVSKLYTDDMPETVETVTTAQTKEKRG						
OTKKEYSIKTTLKELVHKVATSPEDMMQMPPDSYIEMMAQPGENILKNLLEICLT						
ARTKAPDLILEAKETSKLINSLPDIRTICRFAPHGMNIVKVCCHAICCVLRQKKR						
NTVLFHPGATGSKSIIAQIAQAVNGVCYNAANVPFNDCTNNKLIWESAGNFGQ						
QNVQFKAI CSQAIRIDQKGSQKIEPTPVMTNNENITVVRIGCEERPEHTQPTIR						
RMLNIHLTLPQDFGLVDKNEWPMICAWLVKNGQSTWASYCAWKGVYPDWSENAW						
PKVPTPINLLGSGSPFTPKSTPLSONVALTPLASDLEDLALPWSWTENTPVAGTAE						
TQNTGEAGSKACQDGLSFTWSEIEDLRACFAEPLKKDFSEPLND"						
ORIGIN						
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;						
Best Local Similarity 99.9%; Pred. No. 0;						
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;						
Qy	1	ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60			
Db	1	ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60			
Qy	61	AGTAACACGAGAACTGTTCTCATTTGTTTTTAAAAATGAAAAATGTTTCAACTGAATGAAA	120			
Db	61	AGTAACACGAGAACTGTTCTCATTTGTTTTTAAAAATGAAAAATGTTTCAACTGAATGAAA	120			



JOURNAL MEDLINE COMMENT	prototype strain J. Virol. 570, 656-669 (1986) 86115415 Original source text: Minute virus of mice (lymphotropic variant of MMV) DNA, clone pEG222. location/Qualifiers 1. .5085 /organism="Mice minute virus" /mol_type="genomic DNA" /isolate="lymphotropic variant" /db_xref="taxon:10794" /clone="pEG222" 115. .2280 /gene="NS1" 115. .2280 /gene="NS1" /note="putative" /codon_start=1 /product="nonstructural protein" /protein_id="AAA69566.1" /db_xref="GI:825477" /translation="MISGESLQCAKRWAKVYKQMLKSVTYPPFFHSVSDAQKE SNQITMAGNAYSDEVLGTNNLKEKSNQEVFSFVKTEVDQLNGKDIQNMNKKELQOE DELKSLQRAETTDQSGEDMEWSTVDEMTKKQVFIYDSLKKCLFEVLSTKNIAPAE VTWVQHEWGDQGHCHVLIGKDFSAQKQWRRQLNYSRWLVTACNQLTPAE RIKLREIAEDSEWTLTYKHQTKDYKCVLFGNNIAYFLTKKISSTPPRDGY FLASDGMKTNFLKEGERHLVSKLYTDDMRPETVETVTTAQETKGRIGIQTKVESIK TTLKELVHKVTSPPEDMMQPDYSIEMMAQPGENLLKNTLEICILTARTKTAEDL ILEKETSKTUNSLPDTRTCKIFAFHGMNYSVCKHAI CCVLRQGGKRNVLVPHGPA STKESIITAOIAQAVNGCYNAANVPFNDCKNLIWVEAGNFGQOQVFKAIC SGQIRIDQKGGSKQIEPTFVIMTNNITNIVRIGCEERPEHTQPIRDRMLNIHLTH LPGDFGLVDKNWPMICAWLVNGYOSTWASYCAKQKVPDSENWABPKVPTPI LGSARSPTTPKSTPLSONYAITPLASDLEDLALPEWSTENTPVAGTAETQNTGEAGS KACQDQGLSTWSEIEDLRACFGAEPLKDFSEPLNLD" 201. .>2280 /gene="NS1" /product="NS-1 mRNA" 262. .2280 /gene="NS1" /codon_start=1 /product="nonstructural protein" /protein_id="AAA69567.1" /db_xref="GI:332291" /translation="MAGNAYSDEVLGTNNLKEKSNQEVFSFVKTEVDQLNGKDIQNMN KKELOEDELKSLQRAETTDQSGEDMEWSTVDEMTKKQVFIYDSLKKCLFEVL STKNIAPADVTWVQHEWGDQGHCHVLIGKDFSAQKQWRRQLNYSRWLVTACNQLTPAE RIKLREIAEDSEWTLTYKHQTKDYKCVLFGNNIAYFLTKKISSTPPRDGY FLASDGMKTNFLKEGERHLVSKLYTDDMRPETVETVTTAQETKGRIGIQTKVESIK TTLKELVHKVTSPPEDMMQPDYSIEMMAQPGENLLKNTLEICILTARTKTAEDL ILEKETSKTUNSLPDTRTCKIFAFHGMNYSVCKHAI CCVLRQGGKRNVLVPHGPA STKESIITAOIAQAVNGCYNAANVPFNDCKNLIWVEAGNFGQOQVFKAIC SGQIRIDQKGGSKQIEPTFVIMTNNITNIVRIGCEERPEHTQPIRDRMLNIHLTH LPGDFGLVDKNWPMICAWLVNGYOSTWASYCAKQKVPDSENWABPKVPTPI LGSARSPTTPKSTPLSONYAITPLASDLEDLALPEWSTENTPVAGTAETQNTGEAGS KACQDQGLSTWSEIEDLRACFGAEPLKDFSEPLNLD" 203. .4558 /gene="vp" 203. .2281 /gene="vp" /note="major transcription start site" /number=1 2007. .2281 /gene="vp" /note="minor transcription start site" /number=1 2010. .2281 /gene="vp" /note="minor transcription start site" /number=1 2282. .2399 /gene="vp" /note="alternative intron" 2282. .2377	gene  CDS	/gene="vp" /note="alternative intron" join(2287. .2317,2400. .4558) /gene="VP1" join(2287. .2317,2400. .4558) /gene="VP1" /codon_start=1 /protein_id="AAA69569.1" /db_xref="GI:332290" /translation="MAPPAPKRAKRWHPFGYKVLGPGNSLDQGEPTNPNDAAKEHDE AYDQIKSGKNPLYFSAADQRFIDQTKADKWGGKGVHYFFRTKFAFAPKLATDSEP GTSVSRAGKRTTRPAYIFINQARAKKLTSSAAQSSQTSMDGTSQDPGGNAVHSA RVERADGPGGGGGGGVSTGSDYNQTHVFLGDGWVEITATLRLVHLNMP KSENYCRIRVHTNTDTSVKNMAKODDHEQIWTWPSLVANAWGWLQSDWOYICNT MSQLNLVSLDQRIENVLVLTATEQDSSGOAIKIYNNDLTACMVAVDNNILPVTPAA NSMETLGFYPWPKPTIASPRYFECVDRDLSVYENQEGTIEHNMGTGPKMNSQSFITI ENQQITLURTGDFAITGYIFDTNPVLUHTWQNNRQIQGPPLLSTFFPEADTADGTL TAQSGRHGATQMEVNNVSEAIRTPAQVGFQCPHNDFEASRAGPFAAPKVPADVTQGV DREANGSVRYSGKHGHNAAHSGPARYTWDETNFGSGRDRDGFIOASPLVWPPP LNGILTNANPIGTKNDIHPNSVFNSTGYPLTAFSPSPVYPOGOIWKDLDELHKPLRH ITAPFVCKNAPQMLVRLGPNLTDQYDPNGATLSRIYTYGTFFWKGLTMRKLRAN TTWNPVQVSVEDNGNSYMSVTKMLPTATGNMQSVPLITRPARNTY" 2287. .2355 /gene="VP1" /note="ORF1; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69568.1" /db_xref="GI:825478" /translation="MAPPAPKRAKRGKGLRDGLVGY" <2287. .2317 /gene="VP1" /number=2 2318. .2399 /gene="VP1" /note="alternative intron" 2333. .2362 /gene="VP1" /note="ORF3; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69570.1" /db_xref="GI:825479" /translation="MVGWGGINV" 2355. .2399 /gene="VP1" /note="ORF2; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69571.1" /db_xref="GI:825480" /translation="MFNYLFYRPEITWF" 2400. .>4558 /gene="VP1" /number=2 2795. .4558 /gene="VP2" 2795. .4558 /gene="VP2" /codon_start=1 /protein_id="AAA69572.1" /db_xref="GI:332292" /translation="MSDGTSPDGGNAVHSAARADPGGGGGGGGGVSTG GSYDQTHYRFLGDCGWVEITATLRLVHLNMPKSENYCRIRVHTNTDTSVKNMAKOD DHEQIWTWPSLVANAWGWLQSDWOYICNTMSQLNLVSLQRIENVLVLTATDSEP GGQAIKIYNNDLTACMVAVDNNILPVTPAANSMTETLGFYKPKPTIASPRYFPCVD RDLSTYENQEGTIEHNMGTGPKMNSQSFITIENQQITLRTGDEFATGYTFDTNP VKLTHWTQNNRQIQGPPLLSTFFPEADTADGTLTAQSGRHGATQMEVNNVSEAIRTPA QVGFQCPHNDFEASRAGPFAAPKVPADVTQGVDRANGSVRYSGKHGHNAAHSGPA PERYTWDETNGSGRDRDGFIOASPLVWPPP LNGILTNANPIGTKNDIHPNSVFNSTGYPLTAFSPSPVYPOGOIWKDLDELHKPLRH ITAPFVCKNAPQMLVRLGPNLTDQYDPNGATLSRIYTYGTFFWKGLTMRKLRAN TTWNPVQVSVEDNGNSYMSVTKMLPTATGNMQSVPLITRPARNTY" 2282. .2377
-------------------------------	--	-----------------	---

ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948;	Conservative 0; Mismatches 71; Indels 0; Gaps 0;	
Qy	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60	1021 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAAACCAACTTTTCACTGCTGAC 1080
Db	262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACAACCACTGGTTAAAGGAAAAA 321	1282 TTTGACTTGAATTTAGAAAAAGCTGAAACAGCAAACTAAACCAACTTTTCACTGCTGAC 1341
Qy	61 AGTAACACGAGAGTGTCTCATTTGTTTAAATAAGTAATGTTCAACTGAAATGGAATA 120	1081 ACAGAACCTGCGAATTTTTCCTTTTCATGCTGGAATATGTTTAAAGTTTCCCAATGCT 1140
Db	322 AGTAACACGAGAGTGTCTCATTTGTTTAAATAAGTAATGTTCAACTGAAATGGAATA 381	1342 ACAAGAACCTGCAAGATTTTTCCTTTTCATGCTGGAATATGTTTAAAGTTTCCCAATGCT 1401
Qy	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 180	1141 ATTTGCTGTGTTTAAACAGACAAAGGAGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1200
Db	382 GATATCGGATGGAATATTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 441	1402 ATTTGCTGTGTTTAAACAGACAAAGGAGCAAAAGAAATGCTGTTTATTTATTCATGACCA 1461
Qy	181 CGAGGAGCGAAACTACTTGGGACCAAGCGAGACATGGAATGGGNAACCAAGTGGAT 240	1201 GCAGCACAGGCAAAATCTATTTATTCACAAGCCATAGCAACAAGCTTTGGCAATTTGGT 1260
Db	442 CGAGGAGCGAAACTACTTGGGACCAAGCGAGACATGGAATGGGNAATCTACAGTGGAT 501	1462 GCAGTACAGGCAAAATCTATTTATTCACAAGCCATAGCAACAAGCTTTGGTAAATTTGGT 1521
Qy	241 GAAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAAGTGTTTATTTGAA 300	1261 TGCATAATGCGAGCAATGTAAACTTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT 1320
Db	502 GAAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAAGTGTTTATTTGAA 561	1522 TGCATAATGCGAGCTAATGTGAACCTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT 1581
Qy	301 GTGCTTAACCAAGAAATATATTTCTCGTGTGATGTTTAAATTTGGTTGTGCAACATGAATGG 360	1321 TGGGTAGAAAGAGCTGGTAACTTTTGACACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1380
Db	562 GTGCTTAGCACAAAATATAGCTCTGCTGATGTTTACTTTGTTGTGACGATGAATGG 621	1582 TGGGTAGAAAGAGCTGGTAACTTTTGACACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1641
Qy	361 GGAAGACCAAGCTGGCACTGTCATGTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420	1381 GGTCAAACTATTGCGAATTTGATCAAAAAGGAAAGGAGCAAAAAGATTTGAACCAACCA 1440
Db	622 GGAAGACCAAGCTGGCACTGTCATGTAATTTGGAGGCAAGGACTTTAGTCAAGCT 681	1642 GGTCAAACTATTGCGAATTTGATCAAAAAGGAAAGGAGCAAAAAGATTTGAACCAACCA 1701
Qy	421 CAAGGGAATGTTGGAGAGCACTAAATGTTTACTTGGAGCAGATGTTGGTTGTAACAGCC 480	1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db	682 CAAGGGAATGTTGGAGAGCACTAAATGTTTACTTGGAGCAGATGTTGGTTGTAACAGCC 741	1702 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1761
Qy	481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540	1501 GAAACACACTCAACCAATCAGAGACAGATGCTTAAACATTATCTAAACATACATCTTGCCT 1560
Db	742 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 801	1762 GAAACACACTCAACCAATCAGAGACAGATGCTTAAACATTATCTAAACATACATCTTGCCT 1821
Qy	541 GAGTGGTTACTCTTACTTATTAAGCATAGCAAAACCAAAAAGAGCTATACCAAGTGT 600	1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Db	802 GAGTGGTTACTCTTACTTATTAAGCATAGCAAAACCAAAAAGAGCTATACCAAGTGT 861	1822 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1881
Qy	601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGGAAATTAAGCACTAGT 660	1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTCTGATTGG 1680
Db	862 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGGAAATTAAGCACTAGT 921	1882 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTCTGATTGG 1941
Qy	661 CCACCAAGAGACGGAGGCTATTTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720	1681 TCAGAAAACTGGCGGAGCAAAAGTGCCTTCTTATATAATTTACTAGGTTTCGGCAAGC 1740
Db	922 CCACCAAGAGACGGAGGCTATTTTCTAAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 981	1942 TCAGAAAACTGGCGGAGCAAAAGTGCCTTCTTATATAATTTACTAGGTTTCGGCAAGC 2001
Qy	721 AAAGAGGCGAGGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGCAGAAACG 780	1741 TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db	982 AAAGAGGCGGAAACGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGCAGAAACG 1041	2002 TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 2061
Qy	781 GTTGAACACACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAACTTAAATA 840	1801 GCATCGGATCTCGAGCACTGGCTTTAGAGCTTTAGAGCTTGGAGCACCAAACTACTCTGTTGCG 1860
Db	1042 GTTGAACACACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAACTTAAATA 1101	2062 GCATCGGATCTCGAGCACTGGCTTTAGAGCTTTAGAGCTTGGAGCACCAAACTACTCTGTTGCG 2121
Qy	841 GAAGTTTCTATTAATACTACATTTAAAGAGCTGGTGATATAAAGAGTAACTCACCAGAG 900	1861 GGCACTGCGAGAAACCCAGAACTCTGGGAGAGCTGGTTTCCAAAGCCTGCCAAGATGTCAA 1920
Db	1102 GAGTTTCTATTAATACTACATTTAAAGAGCTAGTGATTAAGAGTAACTCACCAGAA 1161	2122 GGCACTGCGAGAAACCCAGAACTCTGGGAGAGCTGGTTTCCAAAGCCTGCCAAGATGTCAA 2181
Qy	901 GACTGGATGATGTCAGGCGACAGTTTACATTGAAATGATGGCTCAACAGGTGGAGAA 960	1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGTTGGTGGGAACCG 1980
Db	1162 GACTGGATGATGTCAGGCGACAGTTTACATTGAAATGATGGCTCAACAGGTGGAGAA 1221	2182 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGTTGGTGGGAACCG 2241
Qy	961 AACCTGCTGAAAAATAGCTAGAGATTTGTACACTAATCTAGCCAGAAACCAAAACAGCA 1020	1981 TTGAAAGAGAGACTTTGAGCGAGCGCTGAACTTGGACTAA 2019
Db	1222 AACCTGCTGAAAAATAGCTAGAGATTTGTACGCTAATCTAGCCAGAAACCAAAACAGCA 1281	2242 TTGAAAGAGAGACTTTGAGCGAGCGCTGAACTTGGACTAA 2280

RESULT 9  
MOU34253  
LOCUS

MOU34253 4764 bp DNA linear VRL 21-AUG-1996

DEFINITION Mouse parvovirus 1b DNA.  
ACCESSION U34253  
VERSION U34253.1 GI:1464793  
KEYWORDS  
SOURCE Mouse parvovirus 1b  
ORGANISM Mouse parvovirus 1b  
REFERENCE 1 (bases 1 to 4764)  
AUTHORS Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.  
TITLE Molecular characterization of newly recognized rodent parvoviruses  
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)  
JOURNAL 96201434  
MEDLINE 8609486  
PUBMED  
REFERENCE 2 (bases 1 to 4764)  
AUTHORS Besselsen, D.G.  
TITLE Direct Submission  
SUBMITTED (17-AUG-1995) David G. Besselsen, Department of  
JOURNAL Veterinary Pathology, University of Missouri-Columbia, W213  
veterinary Medicine Building, Columbia, MO 65211, USA  
LOCATION/Qualifiers  
FEATURES  
source  
1. 4764  
/organism="Mouse parvovirus 1b"  
/mol\_type="genomic DNA"  
/specific\_host="Mus musculus"  
/db\_xref="taxon:42841"  
ORIGIN  
Query Match 94.2%; Score 1902.2; DB 14; Length 4764;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
Qy 1 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 122 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 181  
Qy 61 AGTAACAGGAAGTGTCTCATTTGTTTTTAAATGAAATGTTCAACTGAATGGAATA 120  
Db 182 AGTAACAGGAAGTGTCTCATTTGTTTTTAAACTGAGGATGTTCAACTAAATGGAATA 241  
Qy 121 GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGAGCTGGAATCTTTACAA 180  
Db 242 GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGAGCTGGAATCTTTACAA 301  
Qy 181 CGAGGAGCGGAACTACTTGGGACCAAGGAGGACATGGAATGGGAAACCACTGGAT 240  
Db 302 CGAGGAGCGGAACTACTTGGGACCAAGGAGGACATGGAATGGGAACTACTAGTGGAT 361  
Qy 241 GAAATGACCAAAAAAGCAAGATTCATTTTGAATCTTTGGTTAAAAATGTTTATTTGAA 300  
Db 362 GAAATGACCAAAAAAGCAAGATTCATTTTGAATCTTTGGTTAAAAATGTTTATTTGAA 421  
Qy 301 GTGCTTAACCAAAAGATATATTTCTGGTGATGTTAAATGTTGTCACATGATGAGTGG 360  
Db 422 GTGCTTAGCAAAAGATATAGCTCTCTGATGTTTACTTGGTTTGGCAGCATGATGAGTGG 481  
Qy 361 GGAAGAAGCAAGGCTGGCACTGCACTACTAATTTGGAGGAAGGACTTTAGTCAAGCT 420  
Db 482 GGAAGAAGCAAGGCTGGCACTGCACTACTAATTTGGAGGAAGGACTTTAGTCAAGCT 541  
Qy 421 CAAGGGAATTTGGGAGAGGCACTAAATTTTACTTGGAGGAGATGGTTGGTAAACAGCC 480  
Db 542 CAAGGGAATTTGGGAGAGGCACTAAATTTTACTTGGAGGAGATGGTTGGTAAACAGCC 601  
Qy 481 TGTAAATGCAACTAACACAGCTGAAAGATTAACCTAAGAGAAATAGCAGAGAGCAAT 540  
Db 602 TGTAAATGCAACTAACACAGCTGAAAGATTAACCTAAGAGAAATAGCAGAGAGCAAT 661  
Qy 541 GAGTGGTTACTTACTTACTTATAAGCATAAGCAAAACCAAAAAAGACTTATACCAAGTGT 600  
Db 662 GAGTGGTTACTTACTTACTTATAAGCATAAGCAAAACCAAAAAAGACTTATACCAAGTGT 721  
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT 660

722 GTTCTTTTGGAAACATGATGCTTACTATTTTAAACCAAGAAATAAGCACTAGT 781  
661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAACTACTTTTAA 720  
782 CCACCAAGAGGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAACTACTTTTAA 841  
721 AAAGAAGGCGAGGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780  
842 AAAGAGGCGAACGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 901  
781 GTTGAACCAACAGTAACCACTGCGCAGGAACTAAGCGGCGCAGAAATTCAAATAAAAA 840  
902 GTTGAACCAACAGTAACCACTGCGCAGGAACTAAGCGGCGCAGAAATTCAAATAAAAA 961  
841 GAAGTTTCTTAAATAACTACACTTAAAGAGCTGGTGCATAAAGAGATGAACCTCACCAGAG 900  
962 GAGGTTTCTTAAATAACTACACTTAAAGAGCTGGTGCATAAAGAGATGAACCTCACCAGAA 1021  
901 GACTGGATGATGATGCGCAGCAGACAGTTCATTTGAAATGATGCTCAACAGAGTGGAGAA 960  
1022 GACTGGATGATGATGCGCAGCAGACAGTTCATTTGAAATGATGCTCAACAGAGTGGAGAA 1081  
961 AACTGCTGAAATAATAGCTAGAGATTTGTAACACTTCTAGCCAGAACCAAAAAACGAA 1020  
1082 AACTGCTGAAATAATAGCTAGAGATTTGTAACACTTCTAGCCAGAACCAAAAAACGAA 1141  
1021 TTTGACTTAAATTTAGAAAAAGCTGAAACAGCAACCACTAAACCACTTTTCACTGCCTGAC 1080  
1142 TTTGACTTAAATTTAGAAAAAGCTGAAACAGCAACCACTAAACCACTTTTCACTGCCTGAC 1201  
1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGCAATATGTTTAAAGTTTGCCTATGCT 1140  
1202 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGCAATATGTTTAAAGTTTGCCTATGCT 1261  
1141 ATTGCTGTGTTTTTAAACAGACAGAGGAGCAAAAGAAATGCTGTTTATTTATGAGCA 1200  
1262 ATTGCTGTGTTTTTAAACAGACAGAGGAGCAAAAGAAATGCTGTTTATTTATGAGCA 1321  
1201 GCAGACACAGGCAAACTTATTTGCAAGCCATAGCAACAGCTTGCCTATGTTGTT 1260  
1322 GCAGACACAGGCAAACTTATTTGCAAGCCATAGCAACAGCTTGCCTATGTTGTT 1381  
1261 TGCTATTAATGCAAGCAATGTAACCTTTTCAATTTAATGCTGTACCAACAGAACTTGAT 1320  
1382 TGCTATTAATGCAAGCAATGTAACCTTTTCAATTTAATGCTGTACCAACAGAACTTGAT 1441  
1321 TGGGTAGAGAGAGCTGGTAACTTTTGGACAGCAAGTAACCACTTTTAAAGCCATTTGCTCT 1380  
1442 TGGGTAGAGAGAGCTGGTAACTTTTGGACAGCAAGTAACCACTTTTAAAGCCATTTGCTCT 1501  
1381 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGCAAAACAGATTTCACCAACCA 1440  
1502 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGCAAAACAGATTTCACCAACCA 1561  
1441 GTCATCATGACCAAACTAGAGCAATTCAGTGGTTCAGAAATAGGCTGCGAAGAGAGCA 1500  
1562 GTCATCATGACCAAACTAGAGCAATTCAGTGGTTCAGAAATAGGCTGCGAAGAGAGCA 1621  
1501 GAACACACTCAACCAATCAGAGCAGATGCTTAACTTATCATCTAATACATACCTTGCCT 1560  
1622 GAACACACTCAACCAATCAGAGCAGATGCTTAACTTATCATCTAATACATACCTTGCCT 1681  
1561 GGTGACTTTGTTTGGTTTGCACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620  
1682 GGTGACTTTGTTTGGTTTGCACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1741  
1621 AATGTTTACCAATCTACCATGGCAAGCTACTGTGCTTAAATGGGCAAGTTCTCTGATTGG 1680  
1742 AATGTTTACCAATCTACCATGGCAAGCTACTGTGCTTAAATGGGCAAGTTCTCTGATTGG 1801  
1681 TCAGAAACTGGGCGGAGCCAAAGGTGCAACTCTTATTAATTTACTAGTTTGGCAGCC 1740

Db 1802 ACAGAAACTGGCGGAGCGAGGTGCGACTCCTATATAAATCACTAGTTCGGCAGCG 1861  
Qy 1741 TCACCATTCAGACACCGAAAGTACCCCTCTCAGCAGAACATATGCACTAATCCACTT 1800  
Db 1862 TCACCATTCAGACACCGAAAGTACCGCTCTCAGCAGAACATATGCACTAATCCACTT 1921  
Qy 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTTGCG 1860  
Db 1922 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTTGCG 1981  
Qy 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTCCCAAGATGGTCAA 1920  
Db 1982 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTCCCAAGATGGTCAA 2041  
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGAGATTGAGACGCTGCTTCGGTGGGAAACCG 1980  
Db 2042 CTGAGCCCAACTTGGTCAGAGATCGAGAGAGATTGAGACGCTGCTTCGGTGGGAAACCG 2101  
Qy 1981 TTGAAGAAAGACTTCAGGAGCGCGCTGAACCTTGACTAA 2019  
Db 2102 TTGAAGAAAGACTTCAGGAGCGCGCTGAACCTTGACTAA 2140

## RESULT 10

PAMVMI  
LOCUS Mouse parvovirus minute virus immunosuppressive variant genome (= MWi).  
ACCESSION X02481  
VERSION 1  
KEYWORDS coat protein; genome; origin of replication; overlapping genes; terminal repeat; unidentified reading frame.  
SOURCE Mouse minute virus  
ORGANISM Mice  
REFERENCE 1 (bases 1 to 5087)  
AUTHORS Sahli, R., McMaster, G. K. and Hirt, B.  
TITLE DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice  
JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)  
MEDLINE 85242059  
PUBMED 3855242  
COMMENT For the fibroblast-specific strain (MWmp) sequence see <PAMW2>. The genomes of MWmp and MWmi (immunosuppressive variant) have more than 96% of their sequence in common.  
Data kindly reviewed (18-JUL-1986) by G. McMaster.

## FEATURES

source 1..5087  
/organism="Mice minute virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10794"  
misc\_feature 1..116  
/note="terminal hairpin region"  
promoter 176..183  
/note="TATA box"  
gene 2405..4558  
/gene="VP1"  
CDS <2405..4558  
/gene="VP1"  
/functions="coat protein"  
/codon\_start=1  
/product="VP1 protein"  
/protein\_id="CAB46507.1"  
/db\_xref="GI:5419928"  
/db\_xref="GOA:P07302"  
/db\_xref="Swiss-Prot:P07302"  
/translation="VPPGYKLGPNLSLDQGEPTNPSDAAKSHDBAYDQYIKSGKNP  
YLYPSAADQRFIDQTKADKDWGGKVGHHYFFRTKRAFAPKLATDSEPTGVSRAGRRT  
RPPAYIFINQARAKLTSSAAQSSQTSOTMSDGTSGPDGNGVHSAARVERAADPGGS  
GGGGGGGGVSTGSLVNDQTHYRFLGDGWEITATATRLVHLNMPKSENYCRIRVHN  
TTDTSVKGKNAKDDAHEQIWPMSLVDANAWGLQPSDWQYICNTMSQLNLSLQDE  
IFNVLTKVTEQDSGGQAIKIYNNDLTACMMVAVDNNILPYTPAANSMETLGFYFWK  
PTIASPYRYFYFCVDRDLDSVTYENQEGTIEHNVMGTPKGMNSQFFTIENQOITLLRTG

DEPATCTYYFDTPNPKLTHWTQTNROLGPPPLSTPEADTDAGTLTAQSSRHGATQM  
EVNWSEAIRTPAQVGFQPHNDPEASRAGPFAAPKVPADVTQPPQDREANGSVRSY  
KGWGNWAAHGAPAPRYTWEDEFNFGSRDRDGFIOAPLVLVPPPLGLTLTNANIG  
TKNDIHFSVNSYGLPTAFSHSPVYPOQIWDKELDEHKPLRHITAPFVCKNNAP  
GOMLVRGLNLTQDYPNGATLSRIVTYGTFFWKGLTMRKLRANTTNWPNVTVQSVSE  
DNGNSYMSVTWMLPTATGNMQSVPLITRPVARNY"  
2792..4558  
/gene="VP2"  
CDS <2792..4558  
/gene="VP2"  
/functions="coat protein"  
/codon\_start=1  
/product="VP2 protein"  
/protein\_id="CAB46508.1"  
/db\_xref="GI:5419929"  
/db\_xref="GOA:P07302"  
/db\_xref="Swiss-Prot:P07302"  
/translation="TMSDCTSDPDGNGVHSAARVERAADPGSGGGGGGGVGVSVS  
TGSYDQTHYRFLGDGWEITATATRLVHLNMPKSENYCRIRVHTNTDTSVKGKNAK  
DAHEQIWPMSLVDANAWGLQPSDWQYICNTMSQLNLSLQDEIFNVLTKVTEQD  
SGQAIKIYNNDLTACMMVAVDNNILPYTPAANSMETLGFYFWKPLRHITAPFV  
DRDLSTYENQEGTIEHNVMGTPKGMNSQFFTIENQOITLLRTGDEFATGYFYFDN  
PKLTHWTQTNROLGPPPLSTPEADTDAGTLTAQSSRHGATQMEVNWSEAIRTRP  
AOWFCOPHNDPEASRAGPFAAPKVPADVTQGVDRANGSVRSYCKQHGKNAHGP  
ABERYTWEDEFNFGSRDRDGFIOAPLVLVPPPLGLTLTNANIGTKNDLHFSNVFS  
YGLPTAFSHSPVYPOQIWDKELDEHKPLRHITAPFVCKNNAPGOMLVRGLNLT  
QDYPNGATLSRIVTYGTFFWKGLTMRKLRANTTNWPNVTVQSVSEVDNGNSYMSVT  
KWL  
P7ATGNMQSVPLITRPVARNY"  
4603..4608  
misc\_feature  
/note="polyadenylation signal"  
4821..4826  
misc\_feature  
/note="polyadenylation signal"  
4878..5087  
misc\_feature  
/note="terminal hairpin region"

## ORIGIN

Query Match 94.2%; Score 1902.2; DB 14; Length 5087;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAAACCACTGGTTAAAGGAAAA 321  
Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTAAATAATCAAAATGTTCAACTGAATGAAAA 120  
Db 322 AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAGGATGTTCAACTGAATGAAAA 381  
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGAGACGAGCTGAATCTTTACAA 180  
Db 382 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGAGACGAGCTGAATCTTTACAA 441  
Qy 181 CGAGGAGCGAAACTACTTGGGACCAAGCAGACATGAATGGGAAACACACAGTGGAT 240  
Db 442 CGAGGAGCGAAACTACTTGGGACCAAGCAGACATGAATGGGAAATCTACAGTGGAT 501  
Qy 241 GAAATGACCAAAAAAGCAAGTATTCTTTTGTGTTTAAATAATGTTTATTTGAA 300  
Db 502 GAAATGACCAAAAAAGCAAGTATTCTTTTGTGTTTAAATAATGTTTATTTGAA 561  
Qy 301 GTGCTTAAACAAAAAGTATATTTTCCTGCTGATGTTTAAATTTGTTGCAACATGAATGG 360  
Db 562 GTGCTTAAACAAAAAGTATATTTTCCTGCTGATGTTTAAATTTGTTGCAACATGAATGG 621  
Qy 361 GGAAGAGCAACAGCTGGCAGCTGCCATGACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420  
Db 622 GGAAGAGCAACAGCTGGCAGCTGCCATGACTAATTTGGAGAAAGGACTTTAGTCAAGCT 681  
Qy 421 CAAAGGAAATGGTGGAGAAAGGCAATTAATATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480  
Db 682 CAAAGGAAATGGTGGAGAAAGGCAATTAATATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 741  
Qy 481 TGTAATGTGCAACTAACACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAGACAAT 540



Db 742 TGTAAATGTCAGCTAAACACAGCTCAAGAAATTAACCTAAGAGAAATAGCAGAAGACAGT 801  
Qy 541 GAGTGGGTACTTACTTACTTATAAGCATTAAGCAACCAAAAGAGACTATACCAAGTGT 600  
Db 802 GAGTGGGTACTTACTTACTTATAAGCATTAAGCAACCAAAAGAGACTATACCAAGTGT 861  
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAATAAGCACTAGT 660  
Db 862 GTTCTTTTGGAAATATGATGCTTACTTATTTTAACTTAAAGAAATAAGCACTAGT 921  
Qy 661 CCAACAAGAGACGAGGCTATTTTCTTAGCAGTCACTCTGGCTGGAAACCTTAACCTTTTA 720  
Db 922 CCGCCAGGAGCGAGGCTATTTTCTTAGCAGTCACTCTGGCTGGAAACCTTAACCTTTTA 981  
Qy 721 AAAGAGCGAGCGCCATCTAGTGCAGCAACTATACACTGATGATGATGCGGCCAGAAAG 780  
Db 982 AAAGAGCGAGCGCCATCTAGTGCAGCAACTATACACTGATGATGATGCGGCCAGAAAG 1041  
Qy 781 GTTGAACACACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840  
Db 1042 GTTGAACACACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTTAAAAA 1101  
Qy 841 GAAATTTCTATTAATACTACTTAAGAGCTGTGATTAAGAGTAACCTCAACAGAG 900  
Db 1102 GAGGTTTCTATTAATACTACTTAAGAGCTGTGATTAAGAGTAACCTCAACAGAG 1161  
Qy 901 GACTGGATGATGTCAGCAGCAGATTAATTCAGATGATGCTCAACAGGTGGAGAA 960  
Db 1162 GACTGGATGATGTCAGCAGCAGATTAATTCAGATGATGCTCAACAGGTGGAGAA 1221  
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCAGAAACCAAAACAGCA 1020  
Db 1222 AACCTGCTGAAAAATACGCTAGAGATTTGTACGCTAACTTAGCAGAAACCAAAACAGCA 1281  
Qy 1021 TTTGACTTAATTTAGAAAAAGCTGAAACCGAGGAACTAACCAACTTTTCACTGCTGAC 1080  
Db 1282 TTTGACTTAATTTAGAAAAAGCTGAAACCGAGGAACTAACCAACTTTTCACTGCTGAC 1341  
Qy 1081 ACAAGAACCTGCAAGATTTTGTCTTTTTCATGCTGGAACATATGTTAAAGTTTGCATGCT 1140  
Db 1342 ACAAGAACCTGCAAGATTTTGTCTTTTTCATGCTGGAACATATGTTAAAGTTTGCATGCT 1401  
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTTATTCAGACCA 1200  
Db 1402 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTTATTCAGACCA 1461  
Qy 1201 GCCAGCACAGCAATCTATTTATTCACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260  
Db 1462 GCCAGTACAGCAATCTATTTATTCACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1521  
Qy 1261 TGCTATAATGAGCAATGTAACCTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320  
Db 1522 TGCTATAATGAGCAATGTAACCTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1581  
Qy 1321 TGGGTAGAAGAGCTGTAACCTTTTGGACAGCAAGTAACAGATTTAAAGCAATTTGCTCT 1380  
Db 1582 TGGGTAGAAGAGCTGTAACCTTTTGGACAGCAAGTAACAGATTTAAAGCAATTTGCTCT 1641  
Qy 1381 GGTCAAACTATTGCAATGATCAAAAGGAAAGGAGGAGCAACAGATTTGAACCAACCA 1440  
Db 1642 GGTCAAACTATTGCAATGATCAAAAGGAAAGGAGGAGCAACAGATTTGAACCAACCA 1701  
Qy 1441 GTCAATCATGCCAAATGAGAACATTACAGTGTGATAGATAGCTGCGAAGAAAGACCA 1500  
Db 1702 GTCAATCATGCCAAATGAGAACATTACAGTGTGATAGATAGCTGCGAAGAGAGACCA 1761  
Qy 1501 GAACACACTCAACCAATCAGACAGAGATGCTTAACATTTCACTTAACACATACCTTGCCT 1560  
Db 1762 GAACACACTCAACCAATCAGACAGAGATGCTTAACATTTCACTTAACACATACCTTGCCT 1821  
Qy 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCGCCCATGATTTGCTTGGTTGGTAAAG 1620

Db 1822 GGTGACTTTGGTTGGTTGACAAAGTAATAGGCCCATGATTTGTGCTTGGTTGGTAAAG 1881  
Qy 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAGTTCTCTGATTGG 1880  
Db 1882 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAGTTCTCTGATTGG 1941  
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGGTTCGGACGC 1740  
Db 1942 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGGTTCGGACGC 2001  
Qy 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800  
Db 2002 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 2061  
Qy 1801 GCATCGATCTCGAGACCTGGCTTTTAGAGCCTTTAGAGCCTTTGGAGCACACCAATACTCTGTTGG 1860  
Db 2062 GCATCGATCTCGAGACCTGGCTTTTAGAGCCTTTAGAGCCTTTGGAGCACACCAATACTCTGTTGG 2121  
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Db 2122 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGGTCAA 2181  
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTTCGGTTCGGAAACCG 1980  
Db 2182 CTGAGCCCAACTTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTTCGGTTCGGAAACCG 2241  
Qy 1981 TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db 2242 TTGAAGAGAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2280

## RESULT 11

MPU12469  
LOCUS  
DEFINITION  
Mouse parvovirus 1 complete genome, nonstructural protein 1 (NS1)  
gene, complete cds and capsid protein (VP1) gene, complete cds.  
ACCESSION  
U12469  
VERSION  
U12469.1 GI:525325  
KEYWORDS  
Mouse parvovirus 1  
SOURCE  
ORGANISM  
Mouse parvovirus 1  
REFERENCE  
1 (bases 1 to 5144)  
AUTHORS  
Ball-Goodrich, L.J. and Johnson, E.  
TITLE  
Molecular characterization of a newly recognized mouse parvovirus  
JOURNAL  
J. Virol. 68 (10), 6476-6486 (1994)  
MEDLINE  
9435951  
PUBMED  
8083985  
REFERENCE  
2 (bases 1 to 5144)  
AUTHORS  
Ball-Goodrich, L.J.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (20-JUL-1994) Lisa J. Ball-Goodrich, Section of  
Comparative Medicine, Yale University School of Medicine, New  
Haven, CT 06520-8016, USA

## FEATURES

source  
Location/Qualifiers  
1..5144  
/organism="Mouse parvovirus 1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35340"  
/clone="PRVXB36; plutB/R13-6; prTBg/H3-1"  
265..2283  
/gene="NS1"  
265..2283  
/gene="NS1"  
/codon\_start=1  
/product="nonstructural protein 1"  
/protein\_id="AAA61405.1"  
/db\_xref="GI:525326"  
/translation="MAGNAYSDEVLTGTTNWLKEKSNQVFSFVFKTEDVQLNGKDIGH  
NNYKLEQDELKSLQGAETTWDSDEMEWETVDEMTKKOFTYDSLVRKCLPEVL  
STNTAPADVTWFOHENGKQDGHCHVLIIGKDFSOAGQKWRQLNRYWYRRLVTA  
CNVQLTPABRIKLRIADSESWTLLTYKHQTKDYKCVLFGNMLIAYFLTKKKS  
TSPPRDGGYFUSSDSGWKNTFLUKGERHLVSKLYLTDMPETVETTTVTTAQTKRGRI



QTKVESIKITTLKELVHKRVTSBEDWMMQPDYSIEMMAQPGGNNLLKNLTLEICTLTL  
ARTKTFDLLEKAETSKLNFSLPDRITCKYIFAFHGWNYIKVCHACCCVLNRQGGKR  
NTVLFPGPASTGSIIIAQAAQVAGNVGVCYKAAANVFPFNDCTNKNLIWVEAGNFGQ  
QVNFQKAI CSQGTIRIDQKGGKQIEPTVIMTWNENITVVKIGCEERPEHTQPIRD  
RMLNHLTHLPDGLVDKNWPMICAWLKGNGYOSTWASYCAKRWGKVPDPTENWAE  
PKVPTPINSIGSARSPTTPKSPPLSONYALTPLASDLEDLALPEKSTENTPVAGTAE  
TQNTGAGSKACQDQLSPTWSEIBEDLRCAFGBEPLKXDFSEPLND"  
/gene="vpi"  
join(2290..2320,2403..4561)  
/genes="vpi"  
join(2290..2320,2403..4561)  
/gene="vpi"  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AA061406.1"  
/db\_xref="GI:525327"  
/translation="MAPPAKRAKRGWVPYGGYKYLGPNSLIDQGEPTNPDAAKEHDE  
AYDKYIKSGNPLYLFSADDRIDTQKADGWGKVGHVFFRTKFAFAPRLASSSEP  
GTSVSIAGKRTKPPAHIFINQARAKKRASLAQOORTOTMSDGAPOPSGSAVQSA  
RVBERADGPGSGGGSGGGGVSTGSDYNQTHYFLPSDQWVEITATSTRVHLNMP  
KSENYCRVRHNTNDRTAGNMAKDDAHEQIWTPLSDNSNAGWVFPQSDWQFINCN  
MSHVNLSLDOELFNVIKTVTEQNTGAERAKVIYNNDLTASMMVALDSNNILPYTPT  
DNQETLGFYFWKTPSPRYRYFNCDSLSVTYDQTSIVDTMANASGLSOPFTIE  
NTQRIOLLRTGDSFATCTYFETPEPIKLSHTWOSNQLGQPOITDLPADNENATLV  
TRGDRSGLTOISGSDNVTEATRVBAQVGFPCPDHNEETSRAGPRKVPVVPVAVTQGN  
EHDANGSLRYTDKQGDWGSNNKSEKFTWDALSYDSGRWADRCFINATPTFTSPFAL  
NNLTNSDPIGNKTAIHYQNVFNSYGLPTAFPHPAPIYPOQIOWDKELDLKHPRLHA  
QAPFVCKNNAAGQLLRLAPLTDQYDPSNSTIRVYTGTFPFWKGLTKLAKLRPNA  
TWNPVYQVSAQYQNEENEYSIHKLWLPATATGNMQSIPLLSRPVARNTY"

ORIGIN

Query Match		94.1%; Score 1899; DB 14; Length 5144;
Best Local Similarity		96.3%; Pred. No. 0;
Matches 1944; Conservative		0; Mismatches 75; Indels 0; Gaps 0;
QY	1	ATGCTGGAAATCCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA 60
DB	265	ATGCTGGAAATCCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA 324
QY	61	AGTAACCAAGGAGTGTCTCATTTGTTTTAAAAATGAAAATGTTCAACTGAAATGGAATA 120
DB	325	AGTAACCAAGGAGTGTCTCATTTGTTTTAAAAATGAAAATGTTCAACTGAAATGGAATA 384
QY	121	GATATCGGATGGAATAGTTTACAAAAGAGCTGAGGAGGACGAGCTGAAATCTTTACAA 180
DB	385	GATATCGGATGGAATTAATACAAAAGAGCTGAGGAGGACGAGCTGAAATCTTTACAA 444
QY	181	CGAGGAGCGAAACTCTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240
DB	445	CGAGGAGCGAAACTCTTGGGACCAAGCGAGGACATGGAATGGGAACTACAGTGGAT 504
QY	241	GAAATGACCAAAAAGCAAGTATTTCATTTTGTATTTTGTGTTAAAAAATGTTTATTGAA 300
DB	505	GAAATGACCAAAAAGCAAGTATTTCATTTTGTATTTTGTGTTAAAAAATGTTTATTGAA 564
QY	301	GTGCTTAAACAAAGAAATATTTCTGTTGATGTTAAATTTGTTTGTGCAACATGAATGG 360
DB	565	GTGCTTAGCAAAAGAAATATAGTCTCTGCTGATGTTACTTGGTTTGTGAGCATGAATGG 624
QY	361	GGAAAGACCAAGCTGGCACTGCCATGACTAATTTGGAGGAAGGACTTTTACTCAAGCT 420
DB	625	GGAAAGACCAAGCTGGCACTGCCATGACTAATTTGGAGGAAGGACTTTTACTCAAGCT 684
QY	421	CAAGGGAATATGTTGGAGAGCACTAAATATTTTACTTGGAGCAGATGGTTGGTAAACAGCC 480
DB	685	CAAGGGAATATGTTGGAGAGCACTAAATATTTTACTTGGAGCAGATGGTTGGTAAACAGCC 744
QY	481	TGTAATGTGCAATACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
DB	745	TGTAATGTGCACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAGT 804
QY	541	GAGTGGGTACTTACTTACTTATAGCATAAGCAACCAAAAAGACTATACCAAGTGT 600
DB	805	GAGTGGGTACTTACTTACTTATAAACATTAAGCAAAACCAAAAAGGACTATACCAATGT 864

QY	601	GTTCTTTTGGAAACAATGATTTGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT 660
DB	865	GTTCTTTTGGAAATATGATTTGCTTACTATTTTAAACCAAAAAAATAAGCACTAGT 924
QY	661	CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAATCTTTTA 720
DB	925	CCGCCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAATCTTTTA 984
QY	721	AAAGAAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAAAAG 780
DB	985	AAAGAAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAAAAG 1044
QY	781	GTTGAAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGGGCAGAAATTTCAAACTAAAAA 840
DB	1045	GTTGAAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGGGCAGAAATTTCAAACTAAAAA 1104
QY	841	GAAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900
DB	1105	GAGGTTTCTATTAAAAACCACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAA 1164
QY	901	GACTGATGATGATGCGCAGCAGACAGTTACATTTGAAATGATGCTCAACACAGGTGAGAA 960
DB	1165	GACTGATGATGATGCGCAGCAGACAGTTACATTTGAAATGATGCTCAACACAGGTGAGAA 1224
QY	961	AACCTGCTGAAAAAATACGCTAGAGATTTGTACATACTAAGCTTACCCAGAACCAAAACAGCA 1020
DB	1225	AACCTGCTGAAAAAATACGCTAGAGATTTGTACATACTAAGCTTACCCAGAACCAAAACAGCA 1284
QY	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGAC 1080
DB	1285	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGAC 1344
QY	1081	ACAAGAACCTGCAAGATTTTGTCTTTTCAATGGCTGGAATATCTGTTAAAGTTTCCCATGCT 1140
DB	1345	ACAAGAACCTGCAAGATTTTGTCTTTTCAATGGCTGGAATATCTGTTAAAGTTTCCCATGCT 1404
QY	1141	ATTGCTGTGTTTTTAAACAGACAAAGGAGCAAAAGAAATGCTGTTTTTATTTCAATGACCA 1200
DB	1405	ATTGCTGTGTTTTTAAACAGACAAAGGAGCAAAAGAAATGCTGTTTTTATTTCAATGACCA 1464
QY	1201	GCCAGACAGGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATGTTGT 1260
DB	1465	GCCAGTACAGGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTTGGTAAATGTTGT 1524
QY	1261	TGCTATAATGCAAGCAATGTAAACTTTTCCATTTTAACTGTGACCTACCAACAGAACTTGATT 1320
DB	1525	TGCTATAATGCAAGCAATGTGAACTTTTCCATTTCAATGACTGTATACCAACAGAACTTAATT 1384
QY	1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCTTTGCTCT 1380
DB	1585	TGGGTAGAAAGAGCTGGTAACTTTTGGACCAACAAAGTAAACCAAGTTTAAAGCCTTTGCTCT 1644
QY	1381	GGTCAAACTATTTCGCAATTTGATCAAAAGGAAAGGAGCAACACAGATTGAAACCAACCA 1440
DB	1645	GGTCAAACTATTTCGCAATTTGATCAAAAGGAAAGGAGCAACAGATTGAAACCAACCA 1704
QY	1441	GTCAATCATGACCAACCAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
DB	1705	GTCAATCATGACCAACCAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1764
QY	1501	GAAACACATCAACCAATTCAGACAGAAATGCTTTAACTTATCATCTAAACACATACATTGCCT 1560
DB	1765	GAAACACATCAACCAATTCAGACAGAAATGCTTTAACTTATCATCTAAACACATACATTGCCT 1824
QY	1561	GGTGAATTTGGTTGGTTGACAAAAATGAAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
DB	1825	GGTGAATTTGGTTGGTTGACAAAAATGAAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1684
QY	1621	AATGGTTTACCAATCTACCAATGGCAAGCTACTGTCTAAATGGGGCAAGTTCCTGATTGG 1680
DB	1885	AATGGTTTACCAATCTACCAATGGCAAGCTACTGTCTAAATGGGGCAAGTTCCTGATTGG 1944

```
QY 1681 TCAGAAACTGGCGGAGCCAAAGTGCCTTAAATTTACTAGTTCGGCAGC 1740
Db 1945 ACGGAAACTGGCGGAGCCGAGGTGCGGACTCTATAAATTCATAGTTCGGCAGC 2004
QY 1741 TCACCATTCAGCACACGAAAGTACGCTCTAGCCAGAACATGCACTTCCACTT 1800
Db 2005 TCACCATTCAGCACACGAAAGTACGCTCTAGCCAGAACATGCACTTCCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTGGC 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTGGC 2124
QY 1861 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGTTCCAAAGCTCCCAAGATGSGTCAA 1920
Db 2125 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGTTCCAAAGCTCCCAAGATGSGTCAA 2184
QY 1921 CTGAGCCCACTTGGTTCAGAGATCGAGAGGATTTGAGAGGCTGCTTCGGTGGGAACCG 1980
Db 2185 CTGAGCCCACTTGGTTCAGAGATCGAGAGGATTTGAGAGGCTGCTTCGGTGGGAACCG 2244
QY 1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGACTAA 2019
Db 2245 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGACTAA 2283

RESULT 12
MOU34254
LOCUS
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE
ORGANISM
Mouse parvovirus 1c
Mouse parvovirus 1c
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 4764)
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE
2 (bases 1 to 4764)
Besselsen,D.G.
Direct Submission
TITLE
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
Location/Qualifiers
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 122 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAACAACCAACTGGTTAAAGGAAAA 181
QY 61 AGTAACCAAGAGTGTCTCATTTGTTTTTAAATGAAATGTTCAACTGAATGAAAA 120
Db 182 AGTAACCAAGAGTGTCTCATTTGTTTTTAAATGAAATGTTCAACTGAATGAAAA 241
QY 121 GATATCGGATGGAATAGTTTACAAAAAAGAGCTGCGAGGAGCAGCTGGAATCTTTACAA 180
Db 242 GATATCGGATGGAATANTTACAGNAGGAGCTGCAAGAGCAGGCTAAATCTTTACAA 301
QY 181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240
```

```
Db 302 CGAGGAGCGGAAACTACCTGGGACCAAGCAGGACATGGAATGGGAATCTACAGTGGAT 361
QY 241 GAATGACCAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTGAA 300
Db 362 GAAGTGACCAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTGAA 421
QY 301 GTGCTTAAACACAAAGAATATATTTCTGGTGATGTTAAATTTGGTTTGTGCAACATGAATGG 360
Db 422 GTGCTTAAACACAAAGAATATATTCCTGGTGATGTTAAATTTGGTTTGTGCAATGAATGG 481
QY 361 GGAAGAAGCAAGGCTGGCACTGCCATGTAATAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
Db 482 GGAAGAAGCAAGGCTGGCACTGCCATGTAATAATTTGGAGGCAAGACATTTAGTCAAGCT 541
QY 421 CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGGTAAACAGCC 480
Db 542 CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGGTAAACAGCC 601
QY 481 TGTAAATGTCAACTAACACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAAAGCAAT 540
Db 602 TGTAAATGTCAACTAACACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAAAGCAAT 661
QY 541 GAGTGGTTTACTTCTACTTATTAAGCAATAAGCAAAACCAAAAGACATATACCAAGTGT 600
Db 662 GAATGGTTTACTTCTACTTATTAAGCAATAAGCAAAACCAAAAGACATATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATTCCTACTATTTTAACTTAAAGAAATTAAGCACTAGT 660
Db 722 GTTCTTTTGGCAACATGATTCCTACTATTTTAACTTAAAGAAATTAAGCACTAGT 781
QY 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db 782 CGCCCAAGGAGCGAGGCTATTTCTTAGTACTGCTCTGGCTGGAAAACTAACTTTTAA 841
QY 721 AAAGAAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCCAGAAAAG 780
Db 842 AAAGAAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCCAGAAAAG 901
QY 781 GTTGAACACACAGTACCACTCGCAGGAAACTTAGCGCGGAGAACTTCAAACTTAAAAA 840
Db 902 GTTGAACACACAGTACCACTCGCAGGAAACTTAGCGCGGAGAACTTCAAACTTAAAAA 961
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGTCATAAAAAGAGTAACTCACCAGAG 900
Db 962 GAGGTTTCTATTAAAAACCACTTAAAGAGCTAGTGCTATAGAGAGTAACTCACCAGAA 1021
QY 901 GACTGATGATGATGAGCGCAGACAGTTCATTTGAAATGATGCTCAACAGGTGGAGAA 960
Db 1022 GACTGATGATGATGAGCGCAGACAGTTCATTTGAAATGATGCTCAACAGGTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTCACATAACTCTAGCCAGAACCAAAAACAGCA 1020
Db 1082 AACCTGCTGAAAAATACGCTAGAGATCTGTACATAACTCTAGTAGAAGCAAAAACAGCA 1141
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTGAATTTAGAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCCTGAC 1201
QY 1081 ACAGAAACCTGAGAAATTTTGTCTTTTGTGCTGGAATCTATGTTAAAGTTTGCCTATGT 1140
Db 1202 ACAAGAAACCTGCAAGATCTTTGCTTTTCAATGCTGGAATCTATGTTAAAGTTTGCCTATGT 1261
QY 1141 ATTTGCTGTGTTTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTCATGAGCA 1200
Db 1262 ATTTGCTGTGTTTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTCATGAGCA 1321
QY 1201 GCCAGCACAGGCAATCTATTATTGCAAGCCATAGCACAAAGCAGTTCGGCAATGTTGT 1260
Db 1322 GCCAGTACAGGCAATCCATTATTGCAAGCCATAGCACAGGAGTTCGGTAAATGTTGT 1381
QY 1261 TGTATTAATGAGCCAAATGTTAACTTTTCAATTAATGACTGTACCAACAGAACTTGAT 1320
```

Db 1382 TGTATATGTCAGCAAAATGTGAATCTTCCATTCAATGACTGTACCAACAAGAACTTAATT 1441  
QY 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGATTTAAAGCATTTCCTCT 1380  
Db 1442 TGGGTGGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGATTTAAAGCATTTCCTCT 1501  
QY 1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGGCAAGAGATTTGAAACCAACCA 1440  
Db 1502 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGGCAAGAGATTTGAAACCAACCA 1561  
QY 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500  
Db 1562 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1621  
QY 1501 GAACACATCAACCAATACAGAGACAGATGCTTAAACATTCATCAACATACATTCGCTT 1560  
Db 1622 GAACACATCAACCAATACAGAGACAGATGCTTAAACATTCATCAACATACATTCGCTT 1681  
QY 1561 GGTGACTTTGGTTGGTTGACAAATGAAATGAGTGGCCCATGATTTGCTGGTTGGTAAAG 1620  
Db 1682 GGTGACTTTGGTTGGTTGACAAATGAAATGAGTGGCCCATGATTTGCTGGTTGGTAAAG 1741  
QY 1621 AATGGTTTACCAATCTACATGCGCAAGCTACTGTCTAAATGGGGCAAAAGTTCCTGATTGG 1680  
Db 1742 AATGGTTTACCAATCTACATGCGCAAGCTACTGTCTAAATGGGGCAAAAGTTCCTGATTGG 1801  
QY 1681 TCAGAAAACCTGGCGGAGCCAAAGTGCCAACTCTTAAATTTACTAGGTTCCGCAACGC 1740  
Db 1802 ACGGAAAACCTGGCGGAGCCAAAGTGCCAACTCTTAAATTTACTAGGTTCCGCAACGC 1861  
QY 1741 TCACCAATTCAGACACCGAAAGAGTACGCTCTCAGCCAGAACTATGCACTAATCCACTT 1800  
Db 1862 TCACCAATTCAGACACCGAAAGAGTACGCTCTCAGCCAGAACTATGCACTAATCCACTT 1921  
QY 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAATCTCCTGTTGG 1860  
Db 1922 GCATCGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAATCTCCTGTTGG 1981  
QY 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGTC 1920  
Db 1982 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGTC 2041  
QY 1921 CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGCGTCTCGTGGGGAACCG 1980  
Db 2042 CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAAAGCGTGTTCGGTGGGGAACCG 2101  
QY 1981 TTCAAGAAAGACTTCAGCGAGCCGCTGAACCTGCACTAA 2019  
Db 2102 TTGAAGAGACTTCAGCGAGCCGCTGAACCTGCACTAA 2140

RESULT 13  
HOU34255  
LOCUS  
DEFINITION Hamster parvovirus DNA.  
ACCESSION U34255  
VERSION U34255.1 GI:1464792  
KEYWORDS  
SOURCE Hamster parvovirus  
ORGANISM Hamster parvovirus  
REFERENCE 1 (bases 1 to 4773)  
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.  
TITLE Molecular characterization of newly recognized rodent parvoviruses  
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)  
96201434  
MEDLINE 8609486  
PUBMED  
REFERENCE 2 (bases 1 to 4773)  
AUTHORS Besselsen,D.G.  
TITLE Direct Submission  
Submitted (17-AUG-1995) David G. Besselsen, Department of  
Veterinary Pathology, University of Missouri-Columbia, W213

FEATURES  
source  
Location/Qualifiers  
1. .4773  
/organism="Hamster parvovirus"  
/mol\_type="genomic DNA"  
/specific\_host="Meocricetus auratus"  
/db\_xref="taxon:42843"  
ORIGIN  
Query Match 92.2%; Score 1862.2; DB 14; Length 4773;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60  
Db 122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCACTGGTTAAAGGAAA 181  
QY 61 AGTAACCAAGAGTGTCTCATTTTAAAAATGAAATGTTTCAACTGGAATGAAAA 120  
Db 182 AGTAACCAAGAGTGTCTCATTTTAAAAATGAAATGTTTCACTCAATGGAATA 241  
QY 121 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGAGGACGAGCTGAAATCTTTACAA 180  
Db 242 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGAGGAGAGCTGAAATCTTTACAA 301  
QY 181 CGAGGAGCGGAAACTTACTCTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240  
Db 302 CGAGGAGCGGAAACTTACTCTGGGACCAAGCGAGGACATGGAATGGGAAATCTTCACTGGAT 361  
QY 241 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTTGGTTTAAAAATGTTTATTGAA 300  
Db 362 GAAATGACCAAAAGCAAGTATTTCATTTTGGTCTTCTTACTCTTAAAAATGTTTATTGAA 421  
QY 301 GTGCTTAAACAAAGAAATATATTCTCTGGTGTATTTAAATTTGGTTTGTGCAACATGAATGG 360  
Db 422 GTGCTGAGTACAAAGAAACATAGCACCTAGTGTATTTGTTGTTGTCAGCATGAATGG 481  
QY 361 GGAAGAAGCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTAGTCAAGCT 420  
Db 482 GGAAGAAGCAAGGCTGGCACTGTCTATTAATTTGGAGGCAAGGACTTTAGCAGCT 541  
QY 421 CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGGAGCAGATGTTGTTAAACAGCC 480  
Db 542 CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGGAGCAGATGTTGTTAAACAGCC 601  
QY 481 TGTAAATGTGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540  
Db 602 TGTAGTGTGAGCTATTACAGCTGAAAGAAATTAAGCTGAGAGATAGCGGAGAACCA 661  
QY 541 GAGTGGTTTACTCTACTTACTTATAGCATATAAGCAAAACCAAAAGACTATATACCAAGTGT 600  
Db 662 GAATGGTCACTTTGCTTACTTATAGCATATAAGCAAAACCAAAAGACTATATACCAAGTGT 721  
QY 601 GTTCTTTTGGAAACATGATTTCTTACTATTTTTAACTTAAAGAAATATAGCACTAGT 660  
Db 722 GTTGTCTTTGAAATATAGTGTCTACTACTTTTATCCAAAGAAATATATGCAAGTGT 781  
QY 661 CCACCAAGAGAGCGAGGCTATTCTTAGCAGTACTCTGGCTGGGAAACCTAATCTTTTAA 720  
Db 782 CCACCAAGAGAGCGAGGCTATTCTTAGCAGTACTCTGGCTGGGAAACCTAATCTTTTAA 841  
QY 721 AAAGAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCCAGAAACG 780  
Db 842 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 901  
QY 781 GTTGAACACACAGTAACCACTGGGAGGAACTTAAGCGGCGGAGATTCAACTTAATAA 840  
Db 902 GTTGAACACACAGTAACCACTGGGAGGAACTTAAGCGGCGGAGATTCAACTTAATAA 961  
QY 841 GAAGTTTCTTATTAATAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900  
Db 962 GAGGTCTCTTATTAATAACCACTTAAAGAGCTGGTGCATATAAGAGTAACTCACCAGAA 1021

Qy 901 GACTGGATGATGTCAGCCAGACAGTTCATTCGAAATGATGGCTCAACACAGGTGGAGAA 960  
Db 1022 GACTGGATGATGTCAGCCAGACAGTTCATTCGAAATGATGGCTCAACACAGGTGGAGAA 1081  
Qy 961 AACCTGCTGAAAAATAGCTAGAGATTGTACACTAATCTAGCCAGAAACAAAACAGCA 1020  
Db 1082 AACCTGCTGAAAAATAGCTAGAGATTGTACACTAATCTAGCCAGAAACAAAACAGCA 1141  
Qy 1021 TTTCGACTTAATTTTAGAAAAAGCTGAAACACAGCAAACTTAACCAACTTTTTCACCTGCCTGAC 1080  
Db 1142 TTTCGACTTAATTTTAGAAAAAGCTGAAACACAGCAAACTTAACCAACTTTTTCACCTGCCTGAC 1201  
Qy 1081 ACAAGAACCTGCAGAAATTTTTCCTTTTCATGGCTGGAACATATGTTAAAGTTTGGCATGCT 1140  
Db 1202 ACAAGAACCTGCAGAAATTTTTCCTTTTCATGGCTGGAACATATATTAAGTTTGGCATGCT 1261  
Qy 1141 ATTTGCTGTGTTTAAACAGACAGAGGCGGAAAAGAAATGCTGTGTTTATTTATTCATGACCA 1200  
Db 1262 ATTTGCTGTGTTTAAACAGACAGAGGCGGAAAAGAAATGCTGTGTTTATTTATTCATGACCA 1321  
Qy 1201 GCCAGCAGGCAAACTCTATTATTGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260  
Db 1322 GCCAGCAGGCAAACTCTATTATTGCAAGCCATAGCACAAGCAGTTGGTAAATGTTGGT 1381  
Qy 1261 TGCTATAATGCGCAATGTAAACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320  
Db 1382 TGCTATAATGCGCAATGTAAACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1441  
Qy 1321 TGGGTAGAAGAGCTGTGTAATTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
Db 1442 TGGGTAGAAGAGCTGTGTAATTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1501  
Qy 1381 GGTCAAACTATTCCGCTTGTATCAAAAAGAAAGGAGGAGCAAGATTTGAACCAACACCA 1440  
Db 1502 GGTCAAACTATTCCGCTTGTATCAAAAAGAAAGGAGGAGCAAGATTTGAACCAACACCA 1561  
Qy 1441 GTCATATGACCAACAAATGAGAACTTACAGTGTGTCAGAAATAGGCTCGGAAGAAAGACCA 1500  
Db 1562 GTCATATGACCAACAAATGAGAACTTACAGTGTGTCAGAAATAGGCTCGGAAGAAAGACCA 1621  
Qy 1501 GAACACACTCAACCAATCAGACAGAAATGCTTAAACATTCATCTAACACATACATCTTGCT 1560  
Db 1622 GAACACACTCAACCAATCAGACAGAAATGCTTAAACATTCATCTAACACATACATCTTGCT 1681  
Qy 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTGCTGTTGTTAAAG 1620  
Db 1682 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTGCTGTTGTTAAAG 1741  
Qy 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCTGATTGG 1680  
Db 1742 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCTGATTGG 1801  
Qy 1681 TCAGAAACTCTGGGGAGCCAAAGGTGCAACCTCTTAATAATTTACTAGTTTGGCAAGC 1740  
Db 1802 ACGGAAACTCTGGGGAGCCAAAGGTGCGCACTCTTAATAATTTACTAGTTTGGCAAGC 1861  
Qy 1741 TCACCAATTCAGCACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTTCCACTT 1800  
Db 1862 TCACCAATTCAGCACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTTCCACTT 1921  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGACCTTTAGAGCCTTGGAGCACACCAATATCTCTGTTGG 1860  
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGG 1981  
Qy 1861 GGCATCTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCGCAAGATGTCAC 1920  
Db 1982 GGCATCTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCGCAAGATGTCAC 2041  
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 1980  
Db 2042 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 2101  
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

Db 2102 TTGAAGAGAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2140  
RESULT 14  
MVU34256  
LOCUS  
DEFINITION  
Mice minute virus DNA.  
ACCESSION  
U34256  
VERSION  
U34256.1 GI:1464795  
KEYWORDS  
Mice minute virus  
ORGANISM  
Mice minute virus  
REFERENCE  
1 (bases 1 to 4761)  
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,  
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.  
Molecular characterization of newly recognized rodent parvoviruses  
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)  
MEDLINE  
96201434  
PUBMED  
8609486  
REFERENCE  
2 (bases 1 to 4761)  
Besselsen,D.G.  
Direct Submission  
Submitted (17-AUG-1995) David G. Besselsen, Department of  
Veterinary Pathology, University of Missouri-Columbia, W213  
Veterinary Medicine Building, Columbia, MO 65211, USA  
FEATURES  
source  
1..4761  
/organism="Mice minute virus"  
/mol\_type="genomic DNA"  
/strain="Cutter"  
/specific\_hosts="Mesocricetus auratus"  
/db\_xref="taxon:10794"  
ORIGIN  
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;  
Best Local Similarity 94.9%; Pred. No. 0;  
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 122 ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 181  
Qy 61 AGTAACACAGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
Db 182 AGTAACACAGTGTAGTATCTCATTTGTTTAAAAATGAAGATGTTCAATGAATGAAAA 241  
Qy 121 GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 180  
Db 242 GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTAATAATCTTTACAA 301  
Qy 181 CGAGGAGCGAAACTACTTGGGACCAACGAGGACATGGAATGGGAAACCACTGGAT 240  
Db 302 CGAGGAGCGAAACTACTTGGGACCAACGAGGACATGGAATGGGAACTTTCACTGGAT 361  
Qy 241 GAAATGACCAAAAAAGCAAGTATTTCTTTTGTGTTTAAAAATGTTTATTTGAA 300  
Db 362 GAACTAACCAAAAGCAGATTTTCATTTTGACTCTTTAGTTTAAAAAGTGTATTTTGA 421  
Qy 301 GTGCTTAAACACAAAGAAATATTTCTGCTGATGTTAAATTTGTTGTGCAACATGAATCG 360  
Db 422 GTGCTAAGTACAAAGAAACATAGCTCCTAGTGTATTTAATTTGTTATGTCAGCATGAATCG 481  
Qy 361 GGAAGAACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACATTTAGTCAAGCT 420  
Db 482 GGAAGAACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACATTTAGTCAAGCT 541  
Qy 421 CAAGGAAATGTTGGAGAAAGCAACTAAATGTTTACTGAGCAGATGTTGTTGTAACAGCC 480  
Db 542 CAAGGAAATGTTGGAGAAAGCAGCAATGTTTACTGAGCAGATGTTGTTGTAACAGCC 601  
Qy 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAACTAAGAAATAGCAGAAAGCAAT 540

```
Db 602 TGCAGTGTGAGCTATCACAGCGGAAAGAAATTAAGCTGAGAGAAATAGCGGAAGACCA 661
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGAGCTATACCAAGTGT 600
Db 662 GAATGGGTCACTTGTCTACTTATAGCATAAGCAAAACCAAAAGAGCTATACCAATGT 721
Qy 601 GTTCTTTTGGAAACATGATGTCTTACTATTTTAACTAAAGAAATAAGCACTAGT 660
Db 722 GTTGTCTTTGGAAATATGATGTCTTACTTACTTTTAAACCAAGAAATAATATGCACTAGT 781
Qy 661 CCACCAAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACTTACTTTTAA 720
Db 782 CCACCAAGGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACTTACTTTTAA 841
Qy 721 AAAGAAGGCGAGGCGCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780
Db 842 AAAGAAGGCGAAAGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 901
Qy 781 GTTGAACCAACAGTAACCACTGCGCAGGAAACTAAGCGCGGCAAGAAATTCAAACTAAAAAA 840
Db 902 GTTGAACCAACAGTAACCACTGCGCAGGAAACTAAGCGCGGCAAGAAATTCAAACTAAAAAA 961
Qy 841 GAAGTTTCTATTAAACTACACTTAAGAGCTGGTGCATAAAGAGCTTAACCTCACCAGAG 900
Db 962 GAGGTTTCTATTAAACCAACACTTAAGAGCTGGTGCATAAAGAGAGTAACCTCACCAGAA 1021
Qy 901 GACTGGATGATGACGAGCAGACAGTTACATTTAAATGATGCTCAACAGGCTGGAGAA 960
Db 1022 GACTGGATGATGACGAGCAGACAGTTACATTTAAATGATGCTCAGCCAGGTTGAGAA 1081
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTTAAAAATACGCTAGAGATCTGTACGCTAACTTAGCTAGAACCAAAACAGCC 1141
Qy 1021 TTTGACTTAATTTTAAAGAAAGCTGAAACGAGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTGAATTTTAAAGAAAGCTGAAACGAGCAAACTAACCACTTTTCACTGCGCGAC 1201
Qy 1081 ACAAGAACCTGCAAGATTTTGTCTTTTTCATGGCTGGAACATATGTTAAAGTTTCCCATGCT 1140
Db 1202 ACAAGAACCTGTAAGATTTTGTCTTTTTCATGGCTGGAACATATGTTAAAGTTTCCCATGCT 1261
Qy 1141 ATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATCTCTGTTTATTTATGAGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATCTCTGTTTATTTATGAGACCA 1321
Qy 1201 GCAGACACAGGCAAACTTATTTGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
Db 1322 GCCAGTACAGGCAAACTCCATCAITGCAACAGCCATAGCAACAGCAGTTGGTAATGTTGGT 1381
Qy 1261 TGCTATAATGCGCCATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1382 TGCTATAATGCGCAAAATGGAATTTTCCATTTCAATGACTGCACCAACAAACCACTGATT 1441
Qy 1321 TGGGTAGAAGAGCTGTGTAATCTTTGGACAGCAAGTAAACAGATTTTAAAGCCATTTGCTCT 1380
Db 1442 TGGGTAGAAGAGCTGTGTAATCTTTGGACAGCAAGTAAACCAATTTAAAGCCATTTGCTCT 1501
Qy 1381 GGTCAAACTATTGCAATGTATCAAAAGAAAGGAAAGGCAAGCAACAGATTTGAACCAACCA 1440
Db 1502 GGTCAAACTATTGCAATGTATCAAAAGAAAGGAAAGGCAAGCAACAGATTTGAACCAACCA 1561
Qy 1441 GTCATCATGACCACCAATAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCACCAATAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1621
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAAACATACCTTGCCT 1560
Db 1622 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAAACATACCTTGCCT 1681
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGGCCATGATTTGTGCTTGGTTGGTAAAG 1620
```

```
Db 1682 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1741
Qy 1621 AATGTTTACAAATTAACATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1742 AATGTTTACAAATTAACATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCTGATTGG 1801
Qy 1681 TCAGAAAACCTGGCGGAGCCAAAGGTGCGCAACTCCTATATAAATTTACTAGGTTGCGGACCG 1740
Db 1802 TCAGAAAACCTGGCGGAGCCAAAGGTGCGCAACTCCTATATAAATTTACTAGGTTGCGGACCG 1861
Qy 1741 TCACCAATTCACGACACCGGAAAGTAGTACGCTCTCAGCCAGAACTATATCACTAACTCCACIT 1800
Db 1862 TCACCAATTCACGACACCGGAAAGTAGTACGCTCTCAGCCAGAACTATATCACTAACTCCACIT 1921
Qy 1801 GCATCGATCTCGAGGACCTGCTTTAGAGCCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTGTTGG 1860
Db 1922 GCATCGATCTCGAGGACCTGCTTTAGAGCCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTGTTGG 1981
Qy 1861 GGCATCTGCAAGAACCCAGAACACTCGGGGAAGCTGTTCCAAAGCCTGCCAAAGATGCTCAA 1920
Db 1982 GGCATCTGCAAGAACCCAGAACACTCGGGGAAGCTGTTCCAAAGCCTGCCAAAGATGCTCAA 2041
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 1980
Db 2042 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2140

RESULT 15
LOCUS CQ786765 5121 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 1 from Patent WO2004018689.
ACCESSION CQ786765
VERSION CQ786765.1 GI:45721778
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS 1990, R. and Malerba, M.
TITLE Anti-neoplastic viruses
JOURNAL Patent: WO 2004018689-A 1 04-MAR-2004;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
source Location/Qualifiers
1..5121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Parvovirus H1\VMV (phH1) - Parvovirus H1 with
promoter P4 and left hairpin from VMV"

ORIGIN
Query Match 87.7%; Score 1771; DB 6; Length 5121;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 282 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 341
Qy 61 AGTAAACAGAGAGTGTCTCATTTGTTTTTAAAAATGAAATCTTCAACTGAATGAAAA 120
Db 342 AGTAAACAGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTTCAACTGAATGAAAA 401
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGAGCTGAAATCTTTACAA 180
Db 402 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGAGCTGAAATCTTTACAA 461
Qy 181 CGAGGAGCGGAAACTACTCTTGGGACCAAGCAGAGCATGGAATGGGAAACCAACAGTGGAT 240
```

Db 462 CGAGGACGGAACTACTTGGGACCAAGAGGAGCATGGAATGGGAAACCAACAGTGGAT 521  
Qy 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTAAATAAGTATTTATTTGAA 300  
Db 522 GAAATGACCAAAAGCAAGTATTCATTTTGGTAAATAAGTATTTATTTGAA 581  
Qy 301 GTGCTTAACCAAAAGATATATTTCTGGTGATGTTAATTTGGTTTGGCAACATGAATGG 360  
Db 582 GTGCTTAACCAAAAGATATATTTCTGGTGATGTTAATTTGGTTTGGCAACATGAATGG 641  
Qy 361 GGAAGAACCAAGCTGGCACTGCCATGCTACTAATTTGGAGGAAGGACTTTAGTCAAGCT 420  
Db 642 GGAAGAACCAAGCTGGCACTGCCATGCTACTAATTTGGAGGAAGGACTTTAGTCAAGCT 701  
Qy 421 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480  
Db 702 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 761  
Qy 481 TGTAAATGTCAAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540  
Db 762 TGTAAATGTCAAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 821  
Qy 541 GAGTGGGTACTTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGGACTATACCAAGTGT 600  
Db 822 GAGTGGGTACTTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGGACTATACCAAGTGT 881  
Qy 601 GTTCTTTTGGAAACATGATGTTGCTTACTTATTTTAACTAAAGAAATAAGCACTAGT 660  
Db 882 GTTCTTTTGGAAACATGATGTTGCTTACTTATTTTAACTAAAGAAATAAGCACTAGT 941  
Qy 661 CCACCAAGAGCGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAACCTAACTTTTAA 720  
Db 942 CCACCAAGAGCGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAACCTAACTTTTAA 1001  
Qy 721 AAAGAACGGAGCGCCATCTAGTGAGCAAACTATACCTATGATGACATGCGGCCAGAAAG 780  
Db 1002 AAAGAACGGAGCGCCATCTAGTGAGCAAACTGTTATCTATGATGAGATGAAACCCAGAAAG 1061  
Qy 781 GTTGAACACCAAGTAACCACTCGCAGGAACTAAGCGCGCAGAAATTCAACTTAAATAA 840  
Db 1062 GTCGAGACCAAGTACCACTGACAGAGAGCTAAGCGCGCAGAAATTCAACTTAGAGAG 1121  
Qy 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGATTAAGAGTAACTCACCAAGAG 900  
Db 1122 GAGTCTCGATTAAACCACTCAAGAGTGGTACATTAAGAGTAACTCACCAAGAA 1181  
Qy 901 GACTGGATGATGTCGACGACAGTATTAATTGAATGATGCTCAACCAAGTGGAGAA 960  
Db 1182 GACTGGATGATGTCGACGACAGTATTAATTGAATGATGCTCAACCAAGTGGAGAA 1241  
Qy 961 AACTGTCTGAAATAACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
Db 1242 AACTGTCTTAAATAACCTAGAGATCTGTACACTGACTTAGCAAGAACCAAAACAGCC 1301  
Qy 1021 TTTGACTTAATTTTAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCTGAC 1080  
Db 1302 TTTGACTTGAATCTGAAAAAGCTGAAACCAAGCAAACTAGCCAACTTTTCCATGGCTAGC 1361  
Qy 1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGGGAACTATGTTAAAGTTTGGCAATGCT 1140  
Db 1362 ACCAGAACCTGTGAATCTTTTGTGAGCATGGCTGGAACTATATTAAGTCTGCCATGCC 1421  
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTTATGAGACCA 1200  
Db 1422 ATCTGTTGTGTGATAGACAGAGGCAAAAGGAACTGTGCTCTTTTACAGGACCA 1481  
Qy 1201 GCCAGCACAGCAAACTTATTTATGCAAGCCATAGCAACAGCAAGTGGCAATGTTGGT 1260  
Db 1482 GCCAGCACAGCAAACTTATTTATGCAAGCCATAGCAACAGCAAGTGGTAAATGTTGGT 1541  
Qy 1261 TGCTATTAATGAGCAATGAACTTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320  
Db 1542 TGTACAAATGCTGCCAATGTGAATCTTTTCCATTTTAAATGACTGTACCAACAAAAAATTTGATT 1601

Qy 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1380  
Db 1602 TGGGTGGAAGAGCTGGTAACTTTGGCCAGCAAGTAAACCAATTCAGAGCTATTGTTCT 1661  
Qy 1381 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA 1440  
Db 1662 GGCACAAACCATACGCATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA 1721  
Qy 1441 GTCATCATGACCAACAATGAGAACATTACAGTGGTGCAGATAGCTGCGAGAAAGACCA 1500  
Db 1722 GTTATTATGACCAACCAACGAGAACTTACCGTGGTTAGAAATAGGCTGTGAGAAAGACCA 1781  
Qy 1501 GAAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAAACATACCTTGCCT 1560  
Db 1782 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAAACATACCTTGCCT 1841  
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620  
Db 1842 GGTGACTTTGGTTGGTTGATGAACGAAATGGCTCTGATCTGTGCTTGGTTGGTGAAG 1901  
Qy 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTCTGATTGG 1680  
Db 1902 AATGGTTTACCAATCTACCATGGCTTGTACTGTCTTAAATGGGCAAAAGTTCTCTGATTGG 1961  
Qy 1681 TCAGAAAACTGGCGGAGGCAAAAGGTGCCAACTCTCTATAAATTTACTAGGTTCCGCAAGC 1740  
Db 1962 TCAGAGGACTGGCGGAGGCGGAGCTAGACACTCTCTATAAATTCGCTAGGTTCAATGCGC 2021  
Qy 1741 TCACCAATTCAGACACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACCTCACTT 1800  
Db 2022 TCACCAATCTGACTCCGAGAGTACGCTCTCAGCCAAACTTACGCTCTTACTCCACTT 2081  
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTTAGAGCCTTGGAGCACACCAATACTCTCTGTTGG 1860  
Db 2082 GCATCGACCTTGGGACCTTAGCTCTAGAGCCTTGGAGCACACCAATACTCTCTGTTGG 2141  
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAAGATGGTCAA 1920  
Db 2142 GGCACCTGCAGAACCCAAACACTGGGAGGCTGGTTCCACAGCCTGCCAAAGTGTCAA 2201  
Qy 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAAACCG 1980  
Db 2202 CGGAGCCCAACTGGTCCGAGATCGAGGCGGATTTGAGAGCTTGTCTTCAAGTCAAGAACAG 2261  
Qy 1981 TTCAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db 2262 TTGGAGAGCACTTCAACGAGGAGCTGACCTTGGACTAA 2300

Search completed: January 22, 2005, 21:19:27

Job time : 8547.75 secs





XX The present sequence is a DNA encoding parvovirus non-structure protein 1  
CC (NS1) variant (7394A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoural diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 2019; DB 5; Length 2019;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
DB 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
  
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
DB 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 180  
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 180  
  
QY 181 CGAGGAGCGGAACCTACTTGGGACCAAGCGAGACATGGAATGGGAAACCACTGTGGAT 240  
DB 181 CGAGGAGCGGAACCTACTTGGGACCAAGCGAGACATGGAATGGGAAACCACTGTGGAT 240  
  
QY 241 GAAATGACCAAAAAGCAAGTATTCATTTTGAATCTTTTGGTTAAAAAATGTTTATTGAA 300  
DB 241 GAAATGACCAAAAAGCAAGTATTCATTTTGAATCTTTTGGTTAAAAAATGTTTATTGAA 300  
  
QY 301 GTGCTTAAACAAAGAATATATTCCTGGTGATGTTTAAATGGTTTGGTGAACATGATGG 360  
DB 301 GTGCTTAAACAAAGAATATATTCCTGGTGATGTTTAAATGGTTTGGTGAACATGATGG 360  
  
QY 361 GGAAGACCAAGCTGGCACTGCGATGTAATTAATGGGAAAGGACTTTAGTCAAGCT 420  
DB 361 GGAAGACCAAGCTGGCACTGCGATGTAATTAATGGGAAAGGACTTTAGTCAAGCT 420  
  
QY 421 CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC 480  
DB 421 CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC 480  
  
QY 481 TGTAAATGTCAACTAACCAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 540  
DB 481 TGTAAATGTCAACTAACCAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 540  
  
QY 541 GAGTGGGTTACTCTTACTTACTTAAGCATAAGCAACCAAAAGAACTATACCAAGTGT 600  
DB 541 GAGTGGGTTACTCTTACTTACTTAAGCATAAGCAACCAAAAGAACTATACCAAGTGT 600  
  
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAGT 660  
DB 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAGT 660  
  
QY 661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGAATCTGGCTGGAAAACTAATCTTTTA 720  
DB 661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGAATCTGGCTGGAAAACTAATCTTTTA 720  
  
QY 721 AAAGAAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGATGCGGCGCAGAAACG 780  
DB 721 AAAGAAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGATGCGGCGCAGAAACG 780  
  
QY 781 GTTGAACCCAGTAACCTCGGAGGAACTAAGCGCGGAGAAATTTCAAACTAAAAA 840  
DB 781 GTTGAACCCAGTAACCTCGGAGGAACTAAGCGCGGAGAAATTTCAAACTAAAAA 840  
  
QY 841 GAAGTTTCTATTAATACTACATTAAAGAGCTGGTGCATAAAAAGAGTAACTCACAGAG 900  
DB 841 GAAGTTTCTATTAATACTACATTAAAGAGCTGGTGCATAAAAAGAGTAACTCACAGAG 900

DB 841 GAAGTTTCTATTAATACTACATTAAAGAGCTGGTGCATAAAAAGAGTAACTCACAGAG 900  
QY 901 GACTGGATGATGATGAGCAGACAGTATTAATGAATGATGCTCAACAGAGTGGAGAA 960  
DB 901 GACTGGATGATGATGAGCAGACAGTATTAATGAATGATGCTCAACAGAGTGGAGAA 960  
  
QY 961 AACTGCTGAAAAATAGCGCTAGAGATTTGTACACTAAGTCTTACCCAGAACCAACAGCA 1020  
DB 961 AACTGCTGAAAAATAGCGCTAGAGATTTGTACACTAAGTCTTACCCAGAACCAACAGCA 1020  
  
QY 1021 TTTTGACTTAATTTTACAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTACCTGCCTGAC 1080  
DB 1021 TTTTGACTTAATTTTACAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTACCTGCCTGAC 1080  
  
QY 1081 ACAAGAACCTGCAAGATTTTGTCTTTTCATGGCTGGAATCTATGTTAAAGTTTCCCATGCT 1140  
DB 1081 ACAAGAACCTGCAAGATTTTGTCTTTTCATGGCTGGAATCTATGTTAAAGTTTCCCATGCT 1140  
  
QY 1141 ATTTGCTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATGCTGTTTTTATTTCAATGACCA 1200  
DB 1141 ATTTGCTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATGCTGTTTTTATTTCAATGACCA 1200  
  
QY 1201 GCCAGACAGGCAAACTTATTTGCAAGCCATAGCACAAGCAGTGGCAATGTTGCT 1260  
DB 1201 GCCAGACAGGCAAACTTATTTGCAAGCCATAGCACAAGCAGTGGCAATGTTGCT 1260  
  
QY 1261 TGTATAATGTCAGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTTGATT 1320  
DB 1261 TGTATAATGTCAGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTTGATT 1320  
  
QY 1321 TGGGTAGAGAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTTAAAGCCATTTGCTCT 1380  
DB 1321 TGGGTAGAGAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTTAAAGCCATTTGCTCT 1380  
  
QY 1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAAGATTTGAACCAACCA 1440  
DB 1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAAGATTTGAACCAACCA 1440  
  
QY 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA 1500  
DB 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGTATAGGCTGCGAAGAAAGACCA 1500  
  
QY 1501 GAAACACTCAACCAATCAGAGACAGATGCTTAACTATCATCTAACAATACCTTTGCCT 1560  
DB 1501 GAAACACTCAACCAATCAGAGACAGATGCTTAACTATCATCTAACAATACCTTTGCCT 1560  
  
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAG 1620  
DB 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAG 1620  
  
QY 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCGCAAGTTCTGATTGG 1680  
DB 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCGCAAGTTCTGATTGG 1680  
  
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGTGGCCAACTCTCTATATAATTTACTAGTTTCGACCGC 1740  
DB 1681 TCAGAAAACTGGGCGGAGCCAAAGTGGCCAACTCTCTATATAATTTACTAGTTTCGACCGC 1740  
  
QY 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGCACTATGCATTAATCTCACTTT 1800  
DB 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGCACTATGCATTAATCTCACTTT 1800  
  
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCACCAAAATACTCTGTTGG 1860  
DB 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCACCAAAATACTCTGTTGG 1860  
  
QY 1861 GGCATCTGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGTGTCAA 1920  
DB 1861 GGCATCTGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGTGTCAA 1920  
  
QY 1921 CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGGCTTCGGTGGGAACCG 1980  
DB 1921 CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGGCTTCGGTGGGAACCG 1980



Qy	1981	TTGAGAAAGACTTCAGCGAGCGCGTGAACCTAA	2019
Db	1981	TTGAGAAAGACTTCAGCGAGCGCGTGAACCTAA	2019
RESULT 2			
AAD02797			
ID	AAD02797 standard; DNA; 2019 BP.		
XX	AAD02797;		
AC	(revised)		
DT	06-AUG-2003		
DT	31-MAY-2001 (first entry)		
XX	Parvovirus non-structure protein 1 (NS1) wild-type DNA.		
DE	NS1: non-structure protein 1; cytostatic; gene therapy; toxin; therapy;		
KW	tumoural disease; gene therapy; ds.		
XX	Parvovirus.		
OS			
XX	Key Location/Qualifiers		
FH	1. 2019		
FT	/*tag= a		
FT	/product= "Parvovirus NS1 protein"		
XX	EPI077260-A1.		
PN			
XX	21-FEB-2001.		
XX	13-AUG-1999; 99EP-00115161.		
PF			
XX	13-AUG-1999; 99EP-00115161.		
PR			
XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
PA			
XX	Nueesch J, Rommelaere J;		
PI	WPI; 2001-212717/22.		
XX	P-PSDB; AAY72702.		
DR			
XX	Novel parvovirus non-structure protein variant, useful for treating		
PT	tumoral diseases, has a shifted equilibrium between DNA replication and		
PT	transcription activities, and cytotoxic activity.		
XX	Disclosure; Fig 1; 4lpp; English.		
PS			
XX	The present sequence is a wild type DNA encoding parvovirus non-		
CC	structure protein 1 (NS1). The present invention relates to the variants		
CC	of the parvovirus non-structure protein (NS1) having a shifted		
CC	equilibrium between the DNA replication and transcription activities, and		
CC	the cytotoxicity activity. These variants are useful as toxins for		
CC	treating tumoural diseases. The variant DNAs are useful as vectors for		
CC	gene therapy. (Updated on 06-AUG-2003 to correct OS field.)		
XX			
SQ	Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;		
Query Match 99.9%; Score 2017.4; DB 5; Length 2019;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGGTTAAAGGAAAA	60
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGGTTAAAGGAAAA	60
Qy	61	AGTAACAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120
Db	61	AGTAACAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120
Qy	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA	180
Db	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA	180



```
QY 361 GGAAGACCAAGCTGGCACTGCCATGTACTAATTGGAGGAAGAGCTTTAGTCAAGCT 420
Db 621 GGAAGACCAAGCTGGCACTGCCATGTACTAATTGGAGGAAGAGCTTTAGTCAAGCT 680
QY 421 CAAGGGAATGGTGGAGGCAAGCAATAAATGTTTACTGGAGCAGATGCTGGTTGTAACAGCC 480
Db 681 CAAGGGAATGGTGGAGGCAAGCAATAAATGTTTACTGGAGCAGATGCTGGTTGTAACAGCC 740
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT 540
Db 741 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT 800
QY 541 GAGTGGTACTACTACTTACTATAGCATAAGCAACCAAAAGACTATACCAAGTGT 600
Db 801 GAGTGGTACTACTACTTACTATAGCATAAGCAACCAAAAGACTATACCAAGTGT 860
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAGAAATAAGCACTAGT 660
Db 861 GTTCTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAGAAATAAGCACTAGT 920
QY 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Db 921 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 980
QY 721 AAGAAGGCGAGCGCCATCTAGTGAGCAAACTATATACATGATGACATGCGGCCAGAAACG 780
Db 981 AAGAAGGCGAGCGCCATCTAGTGAGCAAACTATATACATGATGACATGCGGCCAGAAACG 1040
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAACCTAAGCGCGCAGAAATTCAACTAAATAA 840
Db 1041 GTTGAACACACAGTAACCACTGCGCAGGAACCTAAGCGCGCAGAAATTCAACTAAATAA 1100
QY 841 GAAATTTCTATTAAAACTACACTTAAAGAGCTGGTGATAAAAAGAGTAACCTCACCAGAG 900
Db 1101 GAAATTTCTATTAAAACTACACTTAAAGAGCTGGTGATAAAAAGAGTAACCTCACCAGAG 1160
QY 901 GACTGGATGATGACAGCAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA 960
Db 1161 GACTGGATGATGACAGCAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA 1220
QY 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAACCAAAACAGCA 1020
Db 1221 AACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAACCAAAACAGCA 1280
QY 1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1281 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1340
QY 1081 ACAAGAACCTGCGAGAAATTTTGTCTTTTCATGGCTGGAACCTATGTTAAAGTTGCGCATGCT 1140
Db 1341 ACAAGAACCTGCGAGAAATTTTGTCTTTTCATGGCTGGAACCTATGTTAAAGTTGCGCATGCT 1400
QY 1141 ATTTGCTGCTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTTCATGGACCA 1200
Db 1401 ATTTGCTGCTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTTCATGGACCA 1460
QY 1201 GCCAGCAGCGCAATCTATTATTGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT 1260
Db 1461 GCCAGCAGCGCAATCTATTATTGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT 1520
QY 1261 TGCTATAATGAGCCAAATGAAACTTTTCCATTTTAACTGATGTAACCAAGAACTTGATT 1320
Db 1521 TGCTATAATGAGCCAAATGAAACTTTTCCATTTTAACTGATGTAACCAAGAACTTGATT 1580
QY 1321 TGGGTAGAAGAGCTGTTACTTTGGACAGCAGTAACCAAGTTTAAAGCCATTTGCTCT 1380
Db 1581 TGGGTAGAAGAGCTGTTACTTTGGACAGCAGTAACCAAGTTTAAAGCCATTTGCTCT 1640
QY 1381 GGTCAAACTATTCGCTATGATCAAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1440
Db 1641 GGTCAAACTATTCGCTATGATCAAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1700
QY 1441 GTCATCATGACCACAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
```

```
1701 GTCATCATGACCACAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1760
1501 GAAACACATCAACCAATCAGAGACAGAAATGCTTTAAACATTTCATCTAAACACATACCTTGCCT 1560
1761 GAAACACATCAACCAATCAGAGACAGAAATGCTTTAAACATTTCATCTAAACACATACCTTGCCT 1820
1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
1821 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1880
1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTGATTGG 1680
1881 AATGGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTGATTGG 1940
1681 TCAGAAAACTGGCGGAGGACCAAGGTGCCAACTCTCTATAAATTTACTAGGTTCCGCACGC 1740
1941 TCAGAAAACTGGCGGAGGACCAAGGTGCCAACTCTCTATAAATTTACTAGGTTCCGCACGC 2000
1741 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
2001 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060
1801 GCATCGGATCTCGAGACCTGGCTTTAGAGCCTTTGAGGACACACCAATACTCTCTGTGCG 1860
2061 GCATCGGATCTCGAGACCTGGCTTTAGAGCCTTTGAGGACACACCAATACTCTCTGTGCG 2120
1861 GGCACCTGCAGAAACCCAGAACACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
2121 GGCACCTGCAGAAACCCAGAACACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2180
1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAACCG 1980
2181 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAACCG 2240
1981 TTGAAGAAAGACTTCAGCGGCGCTGAACTTGGACTAA 2019
2241 TTGAAGAAAGACTTCAGCGGCGCTGAACTTGGACTAA 2279
```

## RESULT 4

AA02801  
ID AAD02801 standard; DNA; 2019 BP.

XX AAD02801;

XX AC AC

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
tumoural disease; gene therapy; mutant; mutein; variant; ds.

OS Parvovirus.

XX Synthetic.

XX Key Location/Qualifiers

XX CDS 1..2019

FT /tag= a

FT /product= "NS1 variant (T363A) protein"

FT mutation replace(1187, A)

FT /\*tag= b

XX EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99BP-00115161.

XX 13-AUG-1999; 99BP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.



QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATACTCTCTGTTCG 1860  
DB 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATACTCTCTGTTCG 1860  
QY 1861 GGCACTGCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
DB 1861 GGCACTGCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGGTGGCGAACCG 1980  
DB 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGGTGGCGAACCG 1980  
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACCTTGAGACTAA 2019  
DB 1981 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACCTTGAGACTAA 2019

## RESULT 5

AAD02805  
ID AAD02805 standard; DNA; 2019 BP.

XX AAD02805;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX NS1; non-structure protein 1; cytotostatic; gene therapy; toxin; therapy;  
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

FH Key Location/Qualifiers

FT CDS 1..2019

FT /\*tag= a

FT /product= "NS1 variant (T463A) protein"

FT mutation replace(1387, A)

FT /\*tag= b

FT EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J; Rommelaere J;

XX WPI; 2001-212717/22.

XX P-PSDB; AAY72710.

XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.

XX Claim 7; Page 27-30; 41pp; English.

XX The present sequence is a DNA encoding parvovirus non-structure protein 1  
CC (NS1) variant (T463A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoral diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

XX Query Match 99.8%; Score 2015.8; DB 5; Length 2019;

Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAAATCCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60  
DB 1 ATGGCTGGAAATCCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60  
QY 61 AGTAAACAGGAAGTGTTCATTTGTTTAAAAATGAAATGTTTCAACTGAATGGAATA 120  
DB 61 AGTAAACAGGAAGTGTTCATTTGTTTAAAAATGAAATGTTTCAACTGAATGGAATA 120  
QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCGAGGAGCAGCTGGAATCTTTTACAA 180  
DB 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCGAGGAGCAGCTGGAATCTTTTACAA 180  
QY 181 CGAGGAGCGGAAATCTCTTGGGACCAAGAGCTCGAGGAGCAGCTGGAATCTTTTACAA 240  
DB 181 CGAGGAGCGGAAATCTCTTGGGACCAAGAGCTCGAGGAGCAGCTGGAATCTTTTACAA 240  
QY 241 GAAATGACCAAAAAGCAAGTATTTCTTTTGGTTAAATGTTTAAAAATGTTTATTTGAA 300  
DB 241 GAAATGACCAAAAAGCAAGTATTTCTTTTGGTTAAATGTTTAAAAATGTTTATTTGAA 300  
QY 301 GTGCTTTAACACAAAGATATATTTCTCTGATGATTTAAATTTGTTTGGCAACATGAATGG 360  
DB 301 GTGCTTTAACACAAAGATATATTTCTCTGATGATTTAAATTTGTTTGGCAACATGAATGG 360  
QY 361 GGAAGAGCAACAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTTAGTCAAGCT 420  
DB 361 GGAAGAGCAACAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTTAGTCAAGCT 420  
QY 421 CAAAGGAAATGTGGAGAAAGGCAACTAAATTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480  
DB 421 CAAAGGAAATGTGGAGAAAGGCAACTAAATTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480  
QY 481 TGTAAATGTGCACTAACACAGCTGAAGAAATTTAAACTAGAGAAATAGCAGAGACAAT 540  
DB 481 TGTAAATGTGCACTAACACAGCTGAAGAAATTTAAACTAGAGAAATAGCAGAGACAAT 540  
QY 541 GAGTGGTTTACTCTTACTTACTTAAAGCAATAAGCAAAACCAAAAAGACTATACCAAGTGT 600  
DB 541 GAGTGGTTTACTCTTACTTACTTAAAGCAATAAGCAAAACCAAAAAGACTATACCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATTTGCTTACTTATTTTAACTAAAAAGAAAAATAGCAGTGT 660  
DB 601 GTTCTTTTGGAAACATGATTTGCTTACTTATTTTAACTAAAAAGAAAAATAGCAGTGT 660  
QY 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720  
DB 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720  
QY 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
DB 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
QY 781 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAAATAAAA 840  
DB 781 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAAATAAAA 840  
QY 841 GAAAGTTTCTATTTAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900  
DB 841 GAAAGTTTCTATTTAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900  
QY 901 GACTGATGATGATGAGCCAGACAGTTTACATTTGAAATGATGCTCAACACGAGTGGAGAA 960  
DB 901 GACTGATGATGATGAGCCAGACAGTTTACATTTGAAATGATGCTCAACACGAGTGGAGAA 960  
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA 1020  
DB 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA 1020  
QY 1021 TTTGACTCTTAATTTAGAAAAAGCTGGAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080  
DB 1021 TTTGACTCTTAATTTAGAAAAAGCTGGAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080

1021 TTTGACTTAATTTTAGAAAAAGCTGAACCAAGCAAACTAAACAACTTTTCACTGCTGAC 1080  
Db  
1081 ACAGAACCTGCAGATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTTGGCAATGCT 1140  
Qy  
1081 ACAAGAACCTGCAGATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTTGGCAATGCT 1140  
Db  
1141 ATTGCTGTGTTTTAAACAGACAGAGGCAAAAGAAATGCTGTTTTATTATTCATGGACCA 1200  
Qy  
1141 ATTGCTGTGTTTTAAACAGACAGAGGCAAAAGAAATGCTGTTTTATTATTCATGGACCA 1200  
Db  
1201 GCAGACAGGCAATCTATTATTCAGCAAGCCATAGCAGCAAGCAGTTGGCAATGTTGT 1260  
Qy  
1201 GCAGACAGGCAATCTATTATTCAGCAAGCCATAGCAGCAAGCAGTTGGCAATGTTGT 1260  
Db  
1261 TGCTATAATGCAGCCATGTAACCTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320  
Qy  
1261 TGCTATAATGCAGCCATGTAACCTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320  
Db  
1321 TGGGTAGAAGAACTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
Qy  
1321 TGGGTAGAAGAACTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
Db  
1381 GGTCAAACTATTGCGATTGATCAAAAGAAAGGAGGAGCAAGAAAGATTTGAACCAACCA 1440  
Qy  
1381 GGTCAAGCTATTGCGATTGATCAAAAGAAAGGAGGAGCAAGAAAGATTTGAACCAACCA 1440  
Db  
1441 GTCATATGACCAATGAGACATTTACAGTGTGATGATGCTGCGAAGAAAGACCA 1500  
Qy  
1441 GTCATATGACCAATGAGACATTTACAGTGTGATGATGCTGCGAAGAAAGACCA 1500  
Db  
1501 GAACACACTCAACCAATCAGACAGACAGATCTTAAACATTCATCTAACACATACCTTGCCT 1560  
Qy  
1501 GAACACACTCAACCAATCAGACAGACAGATCTTAAACATTCATCTAACACATACCTTGCCT 1560  
Db  
1561 GGTGACTTTGGTTGGTTGACAAAATGAATGAGCCCAATGTTGCTGCTGTTGTTAAAG 1620  
Qy  
1561 GGTGACTTTGGTTGGTTGACAAAATGAATGAGCCCAATGTTGCTGCTGTTGTTAAAG 1620  
Db  
1621 AATGGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG 1680  
Qy  
1621 AATGGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG 1680  
Db  
1681 TCAGAAACTGGGCGAGCCAAAGGTCGCAACTCTTAAATTTACTAGTTTGGCAAGC 1740  
Qy  
1681 TCAGAAACTGGGCGAGCCAAAGGTCGCAACTCTTAAATTTACTAGTTTGGCAAGC 1740  
Db  
1741 TCACCATTCAGCACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Qy  
1741 TCACCATTCAGCACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db  
1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTGCG 1860  
Qy  
1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTGCG 1860  
Db  
1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920  
Qy  
1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920  
Db  
1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTTCCGTTGCGGACCG 1980  
Qy  
1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTTCCGTTGCGGACCG 1980  
Db  
1981 TTGAAGAAAGACTTTCAGGAGCCGCTGAACCTTGACTAA 2019  
Qy  
1981 TTGAAGAAAGACTTTCAGGAGCCGCTGAACCTTGACTAA 2019  
Db

RESULT 6

AAD02799

ID AAD02799 standard; DNA; 2019 BP.

XX

AC

XX

DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.  
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.  
XX Parvovirus.  
OS Synthetic.  
XX Location/Qualifiers  
FH 1..2019  
FT /tag= a  
FT /product= "NS1 variant (S283A) protein"  
FT replace(847, A)  
FT mutation /tag= b  
XX  
PN EP1077260-A1.  
XX  
PD 21-FEB-2001.  
XX  
PP 13-AUG-1999; 99EP-00115161.  
XX  
PR 13-AUG-1999; 99EP-00115161.  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX Nueesch J, Rommelaere J;  
XX  
XX WPI; 2001-212717/22.  
DR P-PSDB; AAY72704.  
XX  
XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.  
XX  
XX Claim 7; Page 11-14; 41pp; English.  
XX  
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein  
CC (NS1) variant (S283A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoral diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;  
Query Match 99.8%; Score 2015.8; DB 5; Length 2019;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGCGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60  
Db 1 ATGCGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60  
Qy 61 AGTAACACAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120  
Db 61 AGTAACACAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120  
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGCTGGAATCTTTTACAA 180  
Db 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGCTGGAATCTTTTACAA 180  
Qy 181 CGAGGAGCGGAAACTACTTGGGACCAAGACATGGAATGGGAAACCAACAGTGGAT 240  
Db 181 CGAGGAGCGGAAACTACTTGGGACCAAGACATGGAATGGGAAACCAACAGTGGAT 240  
Qy 241 GAAATGACCAAAAAAGCAAGATTTCAATTTTGGTTTAAAAAATGTTTATTGAA 300  
Db 241 GAAATGACCAAAAAAGCAAGATTTCAATTTTGGTTTAAAAAATGTTTATTGAA 300

```
QY 301 GTGCTTAACACAAAGATATATTTCTGTGTGATGTTAAATGGTTTGTGCAACATGATGG 360
Db 301 GTGCTTAAACACAAAGAAATATATTTCTGTGTGATGTTAAATGGTTTGTGCAACATGATGG 360
QY 361 GGAAGAGACCAAGCGTGGCACTGCCATGTACTAAATGGAGCAAGAGACTTTTAGTCAAGCT 420
Db 361 GGAAGAGACCAAGCGTGGCACTGCCATGTACTAAATGGAGCAAGAGACTTTTAGTCAAGCT 420
QY 421 CAAGGGAATTTGGTGGAGAGGCAACTAAATTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Db 421 CAAGGGAATTTGGTGGAGAGGCAACTAAATTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAGAAATTAACATAAGACAATAGCAGAGCAAT 540
Db 481 TGTAAATGTGCAACTAACACACAGCTGAAGAAATTAACATAAGAGAAATAGCAGAGCAAT 540
QY 541 GAGTGGTTTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600
Db 541 GAGTGGTTTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTTTAACTAAAGAAATAAAGCACTAGT 660
Db 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTTTAACTAAAGAAATAAAGCACTAGT 660
QY 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAACTAATCTTTTA 720
Db 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAACTAATCTTTTA 720
QY 721 AAGAGAGCGAGCGCCATCTAGTGAGCAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 721 AAGAGAGCGAGCGCCATCTAGTGAGCAACTATACACTGATGACATGCGGCCAGAAACG 780
QY 781 GTTGAACCAACAGTAACCACTGGCGAGGAACTAAGCGCGCAGAAATTCAAACTAAAAAA 840
Db 781 GTTGAACCAACAGTAACCACTGGCGAGGAACTAAGCGCGCAGAAATTCAAACTAAAAAA 840
QY 841 GAAGTTTCTAATTAACACTACCTTAAGAGCTGGTGATTAAGAGTAACCTCACCAGAG 900
Db 841 GAAGTTTCTAATTAACACTACCTTAAGAGCTGGTGATTAAGAGTAACCTCACCAGAG 900
QY 901 GACTGGATGATGACGAGCAGACAGTTACATTGAATGATGCTCAACAGGTTGGAGAA 960
Db 901 GACTGGATGATGACGAGCAGACAGTTACATTGAATGATGCTCAACAGGTTGGAGAA 960
QY 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
Db 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACGAGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACGAGCAAACTAACCACTTTTCACTGCCTGAC 1080
QY 1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTCCATGCT 1140
Db 1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTCCATGCT 1140
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTTATTTTCATGGACCA 1200
Db 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTTATTTTCATGGACCA 1200
QY 1201 GCCAGCACAGGCAAACTATATTGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
Db 1201 GCCAGCACAGGCAAACTATATTGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
QY 1261 TGCTATAATGCGCAATGTAACTTTCCATTTTAACTGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGCTATAATGCGCAATGTAACTTTCCATTTTAACTGACTGTACCAACAGAACTTGATT 1320
QY 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Db 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
QY 1381 GGTCAAACTATTGCAATTGATCAAAAAAGGAAAGGCGACCAACAGATTTGAACCAACCA 1440
```

```
Db 1381 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGCGAGCAACAGATTGAACCAACCA 1440
QY 1441 GTCATCATGACCACCAAAATGAGAACTATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCACCAAAATGAGAACTATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAACTTCACTTAAACATACATACATCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAACTTCACTTAAACATACATACATCTTGCT 1560
QY 1561 GGTGACTTTTGGTTTGTGTCACAAAAATGAATGSCCCATGATTTGTCTTGGTTGGTAAAG 1620
Db 1561 GGTGACTTTTGGTTTGTGTCACAAAAATGAATGSCCCATGATTTGTCTTGGTTGGTAAAG 1620
QY 1621 AATGGTTTACCAATCTACCATGCGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTTACCAATCTACCATGCGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCGCAATCTCTATAAATTTACTAGGTTGCGGACCG 1740
Db 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCGCAATCTCTATAAATTTACTAGGTTGCGGACCG 1740
QY 1741 TCACCAATTCAGACACCGAAAGTAGTCGCTCTCAGCCAGAACTATGCACCTAACCTCCACTT 1800
Db 1741 TCACCAATTCAGACACCGAAAGTAGTCGCTCTCAGCCAGAACTATGCACCTAACCTCCACTT 1800
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGACACCAAAATACTCTCTGTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGACACCAAAATACTCTCTGTTGG 1860
QY 1861 GGCATGCGAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCGCAAGATGGTCAA 1920
Db 1861 GGCATGCGAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCGCAAGATGGTCAA 1920
QY 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGTTGCGGAACCG 1980
Db 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGTTGCGGAACCG 1980
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 7
AAT15311
ID AAT15311 standard; DNA; 5049 BP.
XX
AC AAT15311;
XX
DT 14-OCT-1996 (first entry)
XX
DE Non-attenuated canine parvovirus CPV-39 passage 5 DNA.
XX
KW Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
XX
OS Canine parvovirus.
XX
PN W09614088-A1.
XX
PD 17-MAY-1996.
XX
PF 02-NOV-1995; 95WO-US014207.
XX
PR 08-NOV-1994; 94US-00336345.
XX
PA (CORR ) CORNELL RES FOUND INC.
XX
PI Parriah CR, Gruenberg A, Carmichael LE;
XX
DR WPI; 1996-251556/25.
XX
```



PT Attenuated CPV strains contg. up to 4 mutation (s) relative to control  
PT virus - useful as a veterinary vaccine against CPV disease in animals,  
PT such as wild or domestic dogs.

XX Claim 1; Page 21-24; 42pp; English.

CC This viral DNA is isolated from a non- attenuated CPV. The DNA is  
CC preferably derived from vB1440. The DNA is cloned into a vector which is  
CC used to transfect a host cell. The vector used is preferably pCEM32 or  
CC pCEM52. The host cells to be transfected are selected from Norden  
CC Laboratory feline kidney cells, mink lung cells, Madin-Darby canine  
CC kidney cells or canine A72 cells

XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

Best Local Similarity 71.4%; Pred. No. 4.8e-259;

Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Qy	1	ATGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	60
Db	273	ATGCTGGCAACCACTACTGAGAAAGTTATGAGGGAGTAAATGGTTAAGAAACAT	332
Qy	61	AGTAACGAGAGTGTCTCATTTGTTTTTAAATGAAATGTTCAACTGAATGGAAAA	120
Db	333	GCAGAAATGAAGCAATTTGCTTTGTTTTTAAATGTGACCAACGTCCTCAACTAAATGGAAAG	392
Qy	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTTACAA	180
Db	393	GATGTCGCTGGAAACAATAACCAACCAATTCAAATGAAGCTAACATCTTTAAAT	452
Qy	181	CGAGGAGCGGAACTACTTGGGACCAAGC-----GAGGACATGGAATGGGAACACCA	234
Db	453	AGAGGACACAAACAGCAATGATCAACCCGAAGAAGAAATGGACTGGGAATCGAA	512
Qy	235	GTGATGAATGACCAAAAGCAAGTATTCATTTTTCATTTGTTGTTTAAATGTTTA	294
Db	513	GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGATGCAATTAATTAATAAATGTCCT	572
Qy	295	TTTCAAGTGTCTAACACAAAGAAATATTTCTGTCATGTTAATGTTTGTGCAACAT	354
Db	573	TTTGAAGTCTTTGTTCTTAAATATAGAACCAATGAATGTTGTTGTTTATTCACAT	632
Qy	355	GAATGGGAAAGACCAAGCTGCGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGT	414
Db	633	GAATGGGAAAGATCAAGCTGSCATTTGTCATGTTTACTTCATAGTAAGAACTTACAA	692
Qy	415	CAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGAGCAGATGGTTGGTA	474
Db	693	CAAGCAACTGGTAAATGGCTACGAGACAAATGAATATGTTTGGAGTAGATGGTGGTG	752
Qy	475	ACAGCTGTAATGTGCAACTAACACCAAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAA	534
Db	753	ACTCTTTGTCGGTAACTTAACCAACCACTGAAAGATTAAGCTCAGAGAAATTCAGAA	812
Qy	535	GACAAATGAGTGGGTACTCTACTTACTTATAAGCATAGCAACCAAAAGAGCTATACC	594
Db	813	GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAACAAAGAAAGTATGTT	872
Qy	595	AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATTAAGC	654
Db	873	AAATGTTTCATTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAAAATTTGTC	932
Qy	655	ACTAGTCCCAAGAGAGAGGGCTATTTTCTTAGCAGTGAATCTGCTGCTGAAACCTAAC	714
Db	933	CACAT----GACAAAGAAAGTGGCTATTTTAAAGTACTGATCTCTGTTGAAATTTAAC	989
Qy	715	TTTTTAAAGAGCGGAGCGCACTAGTAGGCAAACTATACACTGATGATGATGCGGCCA	774
Db	990	TTTATGAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAAACAAATGAACCA	1049
Qy	775	GAAACGGTTGAAACCAAGTAAACCACTGCGGAGAACTAAGCGCGGCAGAAATTCAACT	834

Db	1050	GAAACCGTTGAAACCAACGAGTCAGCAGCAGCAAGAAACAAAGCGCGGAGAAATTCAAACT	1109
Qy	835	AAAAAGAGAGCTTTCTATTAAACCTACACTTAAAGAGCTGTCATATAAGAGTAACTCA	894
Db	1110	AAAAAGAGTGTCAATCAATGATGTTCTTCCGGGACTTGGTTAGTAAAGAGTAACTCA	1169
Qy	895	CCAGAGGACTGGATGATGTCAGCAGCAGCAGTGTACATTGAAATGATGGCTCAACAGGT	954
Db	1170	CCTGAAGACTGGATGATGTTTACAACAGATAGTTATATTGAAATGATGGCACAACAGGA	1229
Qy	955	GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCAGAACAAA	1014
Db	1230	GGTGAATACTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTAGCAGAACAAA	1289
Qy	1015	ACAGCATTTGACTTAATTTTAGAAAAAGCTGAAACAGCAACCTAACCAACTTTTCACTG	1074
Db	1290	ACAGCATTTGAAATTAATCTTGAAGAGCAGATATACTAACTAACTAACTTTGATCTT	1349
Qy	1075	CCTGACACAAGAACCTGCGAATTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTGC	1134
Db	1350	GCAAAATTCAGAACATGTCAAAATTTTGTAGAAATGCAAGGATGGAATTTGATTTGTT	1409
Qy	1135	CATGCTATTTGCTGTTTAAACAGACAGGAGGCAAGAAATGCTGTTTATTTTCAT	1194
Db	1410	CACGCTATAGCATGTGTTTAAATAGCAAGGTTGTTAAAGAAATACAGTTCTTTTCAT	1469
Qy	1195	GGACGAGCAGCAGCAGCAAACTTATTATTGCAAGCCATAGCACAAGCAGTTGGCAAT	1254
Db	1470	GGACGAGCAGTACAGGAAATCTATCTGCTCAAGCCATAGCAGCTGTGGTAAAT	1529
Qy	1255	GTTGTTGCTTAAATCAGCCAATTTAACTTTTCCATTTTAATGACTGTACCAACAAGAAC	1314
Db	1530	GTTGTTGTTTAAATCAGCAATGTAAATTTTCCATTTTAACTGCTACCAATAAAAT	1589
Qy	1315	TTGATTTGGTAGAAGAGCTGTAATTTTGGACAGCAAGTAAACAGTTTAAAGCCATT	1374
Db	1590	TTAAATTTGGAATGGAAGAGCTGGTAACTTTGGTCAACAGTTTAACTTAAAGCAATC	1649
Qy	1375	TGCTCTGGTCAAACTATTGCTGTTGATCAAAAGAAAGGAGCAGCAACAGATTTGAACCA	1434
Db	1650	TGTTCTGGACAAACAAATTAGATTGATCAAAAGGTAAGGAAGTAAAGCAATTTGAACCA	1709
Qy	1435	ACACCACTCATCATGACCAACAAATGAGAACATTTACAGTGGTCAAGATAGCTCGAAGAA	1494
Db	1710	ACTCCAGTAAATATGACAACTAATGAAATATAACAATTTGTGAGAATTTGGATGGAAGA	1769
Qy	1495	AGACCAACACACTCAACCAATCAGACAGAGTCTTAACTTAACTTCACTTAAACATACC	1554
Db	1770	AGACCTGAAACATACACAACCAATAGAGACAGAAATGTTGAACATTAAGTTAGTATGTAAG	1829
Qy	1555	TTGCTCTGGTGAATTTGGTTGTTGACAAAATGAAATGGCCCATGATTTGTCTGTTGTTG	1614
Db	1830	CTTCCAGGAGACTTTGGTTGGTTGTTGATAAAGAAATGGCTTTAATATGTCATGGTTA	1889
Qy	1615	GTAAGAAATGGTTACCAATCTCCATGGCAAGCTACTGTGCTTAAATGGGGCAAGTTCTCT	1674
Db	1890	GTTAAACATGGTTTGAATCAACCATGGCTTAACTATACACATCATTTGGGGAAAAAGTACCA	1949
Qy	1675	GATTGCTCAGAAACTGGGGGAGCCAAAGGTCCTTAACTTAAATTTTACTAGTTGTCG	1734
Db	1950	GAATGGATGAAACCTGGGGGAGCTTAAATAAAGAAAGGTTAATAATTCACCGGTTGC	2009
Qy	1735	GCAGCTCACCATTTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACATATGCACTAACT	1794
Db	2010	AAGACT--TAGAGACACAAAGCGCAAGCAATCTCTCAGAGTCAAGCCAAAGTTCTAACT	2066
Qy	1795	CACTTGGCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACAACCAATATCTCT	1854
Db	2067	CTCTGACTCCGACCGCTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATAGCCT	2126
Qy	1855	GTTGGGGGCACTGCAGAAACCCAGAACACTGCGGAGAGCTGTTCCAAAGCCTGCCAAGAT	1914
Db	2127	ATTGCAAGAACTGCAAAATCAACAACTTGGGGCTTACTCA-----CAAGAC	2180



	1915	2181	1975	2241	1974
Qy	GGTCAACTGAGCCCAACTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTCTCGTGGCG		GAACCGTTGAAGAAAGACTTCAGCGA		1974
Db		GTGCAACGCGAGTCCGAGCGTGTCCGAAATAGAGCAGACCTTGAGAGCACTTTTACTTCT		GAACAAATTGGAAGAGGATTTTCGAGA	2240
	1915	2181	1975	2241	

## RESULT 8

AAT15312

ID AAT15312 standard; DNA; 5049 BP.

XX

AC AAT15312;

XX

DT 14-OCT-1996 (first entry) .....

✕

DE Attenuated canine  
vv

**VY**

MY  
MY  
MY  
MY

MM YY  
DD.

## Canine parvovirus

XX  
XX  
:00774047nd 00777533

EH	Key	Location/On
----	-----	-------------

```

...
FT misc feature 59
...

```

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

FT /note= "A.

FT misc feature 97

ET - /★

FT /note= "A,"

FT misc feature 4745

131

### LE

FT misc\_feature 4881

FT /\*tag= d

FT /note= "A G or

	XX	9
		8
		7
		6
		5
		4
		3
		2
		1
		0

PN WO9614088-A1.

[illegible]

FD 17-MAY-1996.

XX  
DE 03 NOV 1965  
03 NOV 1965 1400Z

PF 02-NOV-1995; 93MU-US014207.  
VV

XX  
DD 08-NOV-1994.  
04115-00336315

FK 06-NOV-1994; 2403-00330343.  
YY

PA (COPP ) CORNELI. RES FOUND INC

[illegible]

PT Parrish CR. Gruenberger A. Carmichael J.E.

[illegible]

DR WPI: 1996-251556/25.

**XX**

PT Attenuated CPV strains contg. up to 4 mutation (s) relative to control

PT virus - useful as a veterinary vaccine against CPV disease in animals,

Db	1350	GCAAATCTAGAACATGTC	CAATTTT	TAGATG	CACGGATG	GGAATGG	GATTAAAGTTTCT	1409	XX	Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
Qy	1135	CATGCTATTGCTGCTGTTT	TAACACAG	CAAGGAG	CAAAAGAA	TGCTGTTT	TATTTTCAT	1194	KW	feline panleukopenia virus; mink enteritis virus; infection; ds.
Db	1410	CACGCTATGACATGCTGTTT	TAAATAG	CAAGGTGGT	TAAAGAA	TACAGTCT	CTTTTTCAT	1469	XX	Canine parvovirus; vBI440 (ATCC VR 2489).
Qy	1195	GGACCAAGCCAGCAGACGA	CAATCTATT	TATGCA	CAAGCCAT	AGCAACAG	CTGTGCAAT	1254	FT	Key
Db	1470	GGACCAAGCAAGTACAGAA	AAATCTAT	CTATGCT	CAAGCCAT	AGCAACAG	CTGTGCAAT	1529	FT	Location/Qualifiers
Qy	1255	GTTGGTTGCTATATGATG	CAATGTAA	ACTTTTCC	CAATTTT	TAAGCTGT	TACCAACAGAAC	1314	FT	59
Db	1530	GTTGGTTGTTATATGATG	CAATGTAA	ACTTTTCC	CAATTTT	TAAGCTGT	TACCAATAAAT	1589	FT	/*tag= c
Qy	1315	TTGATTTGGGTAGAGAAG	CTGTAACT	TTTGGAC	AGTAAC	CAGTTTAA	AGCCAT	1374	FT	/note= "base 59 is G in CPV-39 (passage 5)"
Db	1590	TTAAATTTGGATTGAAGA	AGCTGTA	CTTTTGGT	CAACAA	AGTTAAT	CAATTTTAAAGCAAT	1649	FT	97
Qy	1375	TGCTCTGGTCAAACTAT	TGCAATG	ATCAAA	AGGAAAG	GGCAGCA	CAAGATTGAACCA	1434	FT	/*tag= d
Db	1650	TGTTCTGGACAAACAA	TATAGAT	GTCAAA	AGGTAA	AGGAAGTAA	GCMAATTGAACCA	1709	FT	/note= "base 97 is C or T in CPV-39 (passage 5)"
Qy	1435	ACACCACTCATATGAC	CAACAAAT	GAGCA	CAATTA	CACAGT	GTGCTGCGAAGAA	1494	FT	273..2279
Db	1710	ACTCCAGTATATGAC	CAACTAAT	GAAAT	TAAACA	ATTGTG	AGAAATTTGATGTAAGAA	1769	FT	/*tag= a
Qy	1495	AGACCAAGAACACT	CAACCAAT	CAGAC	AGAAAT	GCTTAAC	ATTCTTAACACATACC	1554	FT	/note= "NS1/NS2 coding region"
Db	1770	AGACCTGAACATAC	ACAAACCA	TAAGAC	AGAAAT	GTTGA	CAATTAAGTTAGTATGTAAG	1829	FT	2286..4541
Qy	1555	TTGCTGCTGACATTT	GGTTGGTT	GACAAA	AAATGA	TGGCCCAT	GATTTGTGCTGCTGTTG	1614	FT	/*tag= b
Db	1830	CTTCCAGGACATTT	GGTTGGTT	GTATGA	TAAGAA	GAATGG	CCCTTTAATATGTGTAAGTTA	1889	FT	/note= "VP1/VP2 coding region"
Qy	1615	GTAAGAATGGTTAC	CAATCTAC	CAATG	CAAGCT	ACTTCT	CTAAATGGGCAAAAGTTCTCT	1674	FT	4745
Db	1890	GTTAAACATGTTT	GAATCA	ACCCTG	CTAACT	TATACAT	CATTTGGGGAAAAGTACCA	1949	FT	/*tag= e
Qy	1675	GATTGGTCAGAAAC	TGCGGAG	CCAAAG	GTGCA	ACTCTCT	TATTAATTTACTAGTTCTG	1734	FT	/note= "base 4745 is T in CPV-39 (passage 5)"
Db	1950	GAATGGGATGAAAC	TGCGGAG	CCCTAA	ATACAG	AGGTATA	AAATTCACCAAGTTGC	2009	FT	4881
Qy	1735	GCAGCTCACCATT	TCAGAC	ACCCGAA	AGTAC	CGCTCT	CAGCCAGAACTATGCACTAACT	1794	FT	/*tag= f
Db	2010	AAAGACT---	TAGAC	ACACAG	CGGCA	GCATCT	CAGAGTCAAGACCAAGTTCTAACT	2066	PS	/note= "base 4881 is C in CPV-39 (passage 5)"
Qy	1795	CCACTTCATCGGAT	CTGAG	ACCTG	CTGTTAG	AGCCTT	GAGCCTTGGAGCACACCAATACTCTCT	1854	CC	WO9742972-Al.
Db	2067	CCTCTGACTCGG	ACGTAGT	GGACCT	TTGCACT	GGAACCGT	GGAGTACTCCAGATACGCCT	2126	CC	20-NOV-1997.
Qy	1855	GTTCCGGGCACT	GCAGAA	CCAGAA	CACTG	GGGAAG	CTGTTCCAAAGCCTGCCAAGAT	1914	CC	06-MAY-1997; 97WO-US007584.
Db	2127	ATTGCAGAACT	GCAATCA	CAATCA	AAACCA	CTTGCGG	TTACTCA-----CAAGAC	2180	CC	15-MAY-1996; 96US-00647655.
Qy	1915	GGTCAACTGAC	CCCAACT	TGTCAG	AGATCC	AGAGGAT	TTGAGAGCGTGTTCGGTGGC	1974	CC	(CORR ) CORNELL RES FOUND INC.
Db	2181	GTGCAAGCGAT	CTCGAC	TGTCG	GAATAG	AGCAGC	CTGAGACCTGAGAGCCATCTTTTACTTCT	2240	CC	Parrish CR, Carmichael LE, Gruenberg A;
Qy	1975	GAACCGTTGA	AGAAGACT	CTCAG	CGA	2000			CC	WPI; 1998-008583/01.
Db	2241	GAACAAATTG	GGAAGAT	TTTCG	GAGA	2266			CC	Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
RESULT 9									CC	vaccines for protection against parvovirus and feline pan-leukopenia
AAT88324									CC	virus infections.
ID	AAT88324								CC	Example 8; Page 34-37; 60pp; English.
XX	AC								CC	This DNA sequence comprises an attenuated virus genome derived by serial
XX	XX								CC	passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
DT	17-OCT-2003								CC	39 in NUKF feline kidney host cells. The attenuated virus is designated
DT	21-MAY-1998								CC	(see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
XX	XX								CC	are within the hairpin formed by the 3' terminal palindromic: the mutation
DE	Attenuated canine parvovirus (vBI440) genomic DNA.								CC	at nucleotide 59 introduces an A into a G-C rich region within the tip of
									CC	the hairpin, disrupting the base pairing in one of the 2 small internal
									CC	palindromes within that sequence; the thymine at nucleotide 97 is
									CC	adjacent to the mismatched bubble (flip-flop) sequence within the
									CC	palindrome. The DNA from attenuated CPV strains (see also AAT88321) is
									CC	used for the production of infectious molecular DNA clones, which, in
									CC	turn, can be transfected into cells to generate master stocks of the
									CC	virus. The attenuated viruses can be used in dogs as a vaccine to protect
									CC	against CPV disease, or more generally in cats and minks to protect
									CC	against feline panleukopenia virus and mink enteritis virus. The vaccines
									CC	protect against the currently prevalent CPV-2b type (and all extant
									CC	strains of types 2 and 2a), providing a long term immune response.
									CC	(Updated on 17-OCT-2003 to standardise OS field)
									XX	Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
									XX	Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
									XX	Best Local Similarity 71.4%; Pred. No. 4.8e-269;
									XX	Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
									Qy	1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAGGAAAAA 60

Db 273 ATGTCTGGCAACCAAGTATCTAGGAAGTTATGGAGGAGTAATTTGGTTTAAAGAAACAT 332  
Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTAAATAAGTCAACTGAATGGAATA 120  
Db 333 GCAGAAATGAAGCAATTTCTGTTGTTTAAATGTGACAACTGCAACTTAATGGAAG 392  
Qy 121 GATATCGGATGGAATAGTACAAAAGAGCTGAGAGAGCAAGCTGAAATCTTTACAA 180  
Db 393 GATGTTGCGTGGAACTATACCAACCAATTCAAAATGAAGAGCTAAACATCTTTAAAT 452  
Qy 181 CGAGGAGCGAACTACTCTGGACCAAGC-----GAGGACATGGAATGGGAAACACA 234  
Db 453 AGAGGAGCACAACAGCAATGGATCAACCCGAAGAGAAATGGACTGGGAATCGGAA 512  
Qy 235 GTGGATGAATGACCAAAAAGCAAGTATTCATTTTGGTTCCTTTGTTTAAATAAGTTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGATGCAATTAATTAATAAATGCTT 572  
Qy 295 TTTGAAGTGTCTTAACAAAGAAATATATTTCTGGTGAATGTTAAATTTGGTTGTGCAACAT 354  
Db 573 TTTGAAGTCTTTGTTCTTAAATAATAGAACCAATGAATGTGTTGTTTATTCACAT 632  
Qy 355 GAATGGGAAAGACCAAGCTGCACTGCCATGTACTAATTCGAGGAAGGACTTTAGT 414  
Db 633 GAATGGGAAAGATCAAGGCTGGCATGTGTCATGTTTACTTCATAGTAAGAACTTACAA 692  
Qy 415 CAAGCTCAAGGGAATGGTGGAAAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTA 474  
Db 693 CAAGCAACTGGTAATGGCTACCGACAAATGAATATGTATTTGGAGTAGATGGTTGGTG 752  
Qy 475 ACAGCTGTAAATGTGCACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534  
Db 753 ACTCTTGTTCGGTAAACTTAAACCAACTGAAAGATTAAGCTCAGAGAAATTCAGAA 812  
Qy 535 GACAATGAGTGGTACTTACTTACTTATAAGCATAGCAACCAAAAAGACTATACC 594  
Db 813 GATAGTAATGGGTGACTATATTAACATACAGACATAAGCAAAACAAAAGACTATGTT 872  
Qy 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATAAAGC 654  
Db 873 AAATGGTTCATTTTGGAAATATGATAGCATATTAATTTTAAAGAAATAAATTTGTC 932  
Qy 655 ACTAGTCCACAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAATACTAAC 714  
Db 933 CACAT---GACAAAAGAAAGTGCTATTTTAAAGTACTGATTCCTGGTTGGAATTTAAC 989  
Qy 715 TTTTAAAGAGCGGCGCCATCTAGTGAGCAAACTATACACTGATGATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGNAACCA 1049  
Qy 775 GAAACGGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 834  
Db 1050 GAAACGGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 1109  
Qy 835 AAAAAAGAGTGTCTATTAATAACTACATTTAAAGAGCTGGTGATTAAGAGTAACTCA 894  
Db 1110 AAAAAAGAGTGTCAATCAATGCTACTTTGCGGCACTTGGTTAGTAAAGAGTAAACATCA 1169  
Qy 895 CCAGAGGACTGGATGATGAGCGACAGCAGTTACATTTGAATGATGGCTCAACAGGT 954  
Db 1170 CTTGAAGACTGGATGATGTTTAAACAGATAGTATATTTGAATGATGGCAGCAACAGGA 1229  
Qy 955 GGAGAAACCTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCGCAACCA 1014  
Db 1230 GGTGAAATCTTTTAAATAATACATTTGAATTTGTACTTTGACTTTAGCAAGAACAA 1289  
Qy 1015 ACAGCAATTTGACTTAATTTTAAAGAGCTGAAACCAAGCACTAACCACTTTTCACTG 1074  
Db 1290 ACAGCAATTTGAAATTAATTTGAAAGAGCAGATAACTAACTAACTAACTTTGATCTT 1349  
Qy 1075 CTTGACACAAGAACCTGCGAATTTTGTCTTTTATGCTGGCACTATGTTAAAGTTTGC 1134

Db 1350 GCAAAATCTAGAAACATGTCMAATTTTTAGAAATGCAAGGATGGAATGGAATTAAGTTTGT 1409  
Qy 1135 CATGCTATTTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTTTATTCAT 1194  
Db 1410 CACGCTATAGCATGTGTTTTTAAATAGCAAGGTGGTAAAGAAATACAGTCTCTTTTTCAT 1469  
Qy 1195 GACACGAGCAGCAGCAGCAATCTATTTATTTGCAAGCCCATAGCAGCAGTGTGGCAAT 1254  
Db 1470 GGACACGAGCAAGTACAGGAAATCTATCATTTGCTCAAGCCCATAGCAAGCTGTGGTAAT 1529  
Qy 1255 GTTGGTGTCTATTAATCAGCCCAATGTAACTTTTCCATTTTAACTGTGTACCAACAAGAC 1314  
Db 1530 GTTGGTGTATTAATGCGCAATGTAAATTTTCCATTTTAACTGTGTACCAATTAATAT 1589  
Qy 1315 TTGAATTTGGTGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATT 1374  
Db 1590 TTAATTTGGATTGAAGAAGCTGGTAACTTTTGGTCAACAAGTTAATCAATTTAAAGCAATC 1649  
Qy 1375 TGCTCTGGTCAAACTATTTGCAATTTGATCAAAAAGAAAGGCAAGCAACAGATTTGAACCA 1434  
Db 1650 TGTCTCTGGCAAACTAATTAAGATTTGATCAAAAAGGTAAGGAAATTTGAACCA 1709  
Qy 1435 ACACCAAGTCAATCATGACCAACAAATGAGAACATTTACAGTGTGAGATAGGCTCGAAGAA 1494  
Db 1710 ACTCCAGTAAATTAAGCAACTAATGAATAATTAACAATTTGAGAAATTTGAGATGAAAGAA 1769  
Qy 1495 AGACCAAGAACACACTCAACCAATTCAGAGACAGAAATGCTTAAACATTTCACTTAACACATACC 1554  
Db 1770 AGACCTGAACATACACCAACCAATTAAGACAGAGAAATTTGAACATTTAGTTAGTATGAAG 1829  
Qy 1555 TTGCTCTGGTGAATTTGGTTGGTTGACAAAATTAAGATGGCCCATGATTTGCTGTGGTTG 1614  
Db 1830 CTTCCAGGAGACTTTGTTGGTTGGTAAAGAAATGGCCCTTTAAATATGTCATGCTGTTA 1889  
Qy 1615 GTAAAGATTTGGTTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGCAAGTTCTCT 1674  
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACATCATTTGGGGAAGAAATACCA 1949  
Qy 1675 GATTGCTCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAAATTTACTAGGTTGC 1734  
Db 1950 GAATGGATGAAACCTGGGCGGAGCTTAAATACAAAGAGGTATAAATTCACCAGGTTGC 2009  
Qy 1735 GCAGCTCACCATTACAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794  
Db 2010 AAAGACT---TAGAGACACAAGCGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT 2066  
Qy 1795 CCACTTTGATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGAGCACACCAATACTCCT 1854  
Db 2067 CTTCTGACTCCGACGCTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCT 2126  
Qy 1855 GTTGGCGGCACTGCAAGAAACCCAGAAACACTGGGGAAAGCTGGTTCCAAAGCTGCGCAAGAT 1914  
Db 2127 ATTGCAAGAACTGCAATCAACAATCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180  
Qy 1915 GGTCAACTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGGCTGCTCGGTGCG 1974  
Db 2181 GTGCAAGCAGTCCGAGCTGGTCCGAAATAGAGCGAGACCTGAGAGCCACTTTTACTTCT 2240  
Qy 1975 GAACCGTTGAAGAAAGACTTTCAGCGA 2000  
Db 2241 GAACAATTTGAAGAAGATTTTCGAGA 2266

RESULT 10  
AAT88321  
ID AAT88321 standard; DNA; 5049 BP.  
XX  
AAT88321;  
AC  
XX  
21-MAY-1998 (first entry)  
XX  
Attenuated canine parvovirus genomic DNA.  
DE  
XX



```
Db 1230 GGTGAAATCTTTAAAAATACACTTGAAATTTGTACTTTGACTTTAGCAAGACAAAA 1289
Qy 1015 ACAGCATTTCACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACAACTTTTCACTG 1074
Db 1290 ACAGCATTTGAATTAATTAATTTAGAAAAAGCAGATAATACTAAACTAACTAACTTTGATCTT 1349
Qy 1075 CTTGACACAGAACCTCGAGATTTTTCCTTTTTCATGCTGGTGAACATAATGTTAAAGTTTGC 1134
Db 1350 GCAAAATTTCTAGACATGTCAAAATTTTAGAATGACGAGTGGAAATTCGATTTAAAGTTTGT 1409
Qy 1135 CATGCTATTTGCTGTTTAAACAGACAGAGGCAAAAGAAATGCTGTTTATTTATTTTCAAT 1194
Db 1410 CACGCTATAGCATGTGTTTAAATAGACAGAGGTGTTAAAGAAATACAGTCTTTTTCAT 1469
Qy 1195 GGACCAAGCAGCAGCAAACTTATTTATGCAAGCCATAGACAAAGCAGTTGGCAAT 1254
Db 1470 GGACCAAGCAGCAAACTTATCATTTGCTCAAGCCATAGACAAAGCTGTGGTAAAT 1529
Qy 1255 GTTGGTGTCTATATGAGCAAACTGTAACCTTTCCATTTTAACTGCTACCAAGAAC 1314
Db 1530 GTTGGTGTCTATATGAGCAAACTGTAATTTTCCATTTTAACTGCTACCAATAAAAAT 1589
Qy 1315 TTGATTTGGGTAGAAAGCTGTGTAATTTTGGACAGCAAGTAAACCACTTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATGGAAGCTGTGTAATTTTGGTCAACAGTAACTAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTTGCAATTTGATCAAAAAGGAAAGCAGCAAAAGATTTGAACCA 1434
Db 1650 TGTCTGACAAACAAATAGAAATTTGATCAAAAAGGTAAGGAAGTAAAGCAATTTGAACCA 1709
Qy 1435 ACACAGTATCATGACACAAATAGAAACATTTACAGTGTGAGAAATAGGCTGGGAAGAA 1494
Db 1710 ACTCCAGTAATATGACAACTAATGAAATTTTAACAATTTGTGAGAAATTTGATGTAAGAA 1769
Qy 1495 AGACCAAGACACACTCAACCAATCAGACAGAGATGCTTAACTTCACTTAAACACATACC 1554
Db 1770 AGACCTGACATCACAAACCAATAGAGACAGAAATGTTGAACATTAAGTATGATGTAAG 1829
Qy 1555 TTGCTGTGACTTTGTTGTTGTTGACAAAAATGAATGGCCCATGATTTGTGCTGTGGTG 1614
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGTTGATAAAGAAATGGCTTTTAAATATGTGCATGTTA 1889
Qy 1615 GTAAAGAAATGTTACCAATATACATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACATCATTTGGGGAAGATACCA 1949
Qy 1675 GATTGTCAGAAACTGGGCGGAGCCAAAGTGCCCACTCTCTATAAATTTTACTAGTTGCG 1734
Db 1950 GAATGGGATGAAAACTGGGCGGAGCCCTAAATACAGAAAGTATAAATTCACCAAGTTGC 2009
Qy 1735 GCAGCTCACCATTTCCAGCACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACATACT 1794
Db 2010 AAAGACT--TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGCAAGTTCTTAAT 2066
Qy 1795 CCACTTGCAATCGATCTGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCT 1854
Db 2067 CTTCTGACTCGGACGTAGTGGACCTTTCACCTGGAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGGCACTGAGAAACCCAGAACCTGGGGAAGCTGTTCCAAAGCTGCCAAGAT 1914
Db 2127 ATTGAGAAATGGAATTAACAAATCAACCAACTTGGCGTTACTCA-----CAAGAGC 2180
Qy 1915 GGTCAACTGAGCCCAACTTTGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCG 1974
Db 2181 GTGCAAGGAGTCCGACGTGTGTCGAATAGAGGACGACCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAGAAAGAACTTCAGCGA 2000
Db 2241 GAACAAATTGGAAGAAAGATTTTCGAGA 2266
```

RESULT 11  
AAT88320

```
ID XX AAT88320 standard; DNA; 5049 BP.
XX AC AAT88320;
XX XX 17-OCT-2003 (revised)
DT 21-MAY-1998 (first entry)
XX XX Canine parvovirus 39 passage #5 (wild-type).
DE XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX XX Canine parvovirus; type 2b isolate 39.
XX XX Location/Qualifiers
FH CDS 273..2279
FT /*tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /*tag= b
FT /note= "VP1/VP2 coding region"
XX W09742972-A1.
XX PN 20-NOV-1997.
XX XX 06-MAY-1997; 97WO-US007584.
XX XX 15-MAY-1996; 96US-00647655.
XX XX (CORR ) CORNELL RES FOUND INC.
XX PI Parriah CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
PT vaccines for protection against parvovirus and feline pan-leukopenia
PT virus infections.
XX XX Example 8; Page 37-40; 60pp; English.
XX XX This DNA sequence comprises the genome of virulent canine parvovirus type
CC 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NURK feline
CC kidney host cells. Further passaging has yielded attenuated virus vBI440
CC (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
CC (see AAT88321) derived from the 65th passage (ATCC 2528). These
CC respectively contain 4 and 6 mutations in comparison to the virulent 5th
CC passage virus. The DNA from attenuated CPV-2b strains is used for the
CC production of infectious molecular DNA clones, which, in turn, can be
CC transduced into cells to generate master stocks of the virus. The
CC attenuated viruses can be used in dogs as a vaccine to protect against
CC CPV disease, or more generally in cats and minks to protect against
CC feline panleukopenia virus and mink enteritis virus. The vaccines protect
CC against the currently prevalent CPV-2b type (and all extant strains of
CC types 2 and 2a), and provide a long term immune response. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;
Best Local Similarity 71.3%; Pred. No. 1.3e-268;
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;
Qy 1 ATGCTCGGAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGTTAAAGGAAAA 60
Db 273 ATGCTCGCAACCACTGATGAGAAAGTTATGGAGGAGTAAATTTGTTAAAGAACAT 332
Qy 61 AGTAAACCAAGGAAGTGTCTCATTTGTTTAAATAAATGAAATGTTCAACTGAATGGA 120
Db 333 GCAGAAATGAAGCATTTTCGTTGTTTAAATGTCACACGTCACCACTAAATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGTAATCTTTTACA 180
```





Db 1567 CCAGGTGATTTGGACTTTTAGAAGAACTGTAATGCCACTAATATGTGCTTGGTGGTA 1626  
 QY 1618 AAGAATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAGTTCTGTAT 1677  
 Db 1627 AAGAAAGGTTACCAAGCAACAATGGCTAGCTATATGTCATCATTTGGGGAATGTACCTGAT 1686  
 QY 1678 TGGTCAGAAACTGGGCGAGCCAAAGGTGSCCAACTCTTATAAATTTACTAGTTTCGGCA 1737  
 Db 1687 TGGTCAGAAATAATGGAGGAGCCAAATAATGTCATTTCCCAATAATACACCA---ACAGAC 1743  
 QY 1738 CGCTCACCATTTCAGCACCGAAGAAAGTACGCCTCTCAGCAGAACTATGCACTAATCCA 1797  
 Db 1744 TCTCAGATTTCAGATCAGTGAAACTTGCAGCGAGACATCACTACGAGCAACTCCA 1803  
 QY 1798 CTTGCAATCGATCTCGA---GGACCTGGCTTTAGAGCCTTGGAGCACCACTAATCTCC 1853  
 Db 1804 ATACAGGAGGACCTGGATTAGCTTTAGCCTTGGAGCGGTGGAGCGAGCCAAACACCC 1862

RESULT 13  
 ABQ95626  
 ID ABQ95626 standard; DNA; 374 BP.  
 AC ABQ95626;  
 XX  
 XX  
 DT 28-OCT-2002 (first entry)  
 XX  
 XX  
 XX Tumour suppression-related oligonucleotide #1277.  
 XX  
 XX Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
 KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
 KW viral infection; cell degeneration disease; neurodegeneration; ds;  
 KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.  
 XX

OS Homo sapiens.  
 XX  
 XX FR2819824-A1.  
 XX  
 XX 26-JUL-2002.  
 XX  
 XX 23-JAN-2001; 2001FR-00000899.  
 XX  
 XX 23-JAN-2001; 2001FR-00000899.  
 XX  
 XX (MOLE-) MOLECULAR ENGINES LAB SA.  
 XX  
 XX Telerman A, Amson R, Tuijnder M, Susini L;  
 XX  
 XX WPI; 2002-610803/66.  
 XX  
 XX New nucleic acid implicated e.g. in tumor suppression, useful for  
 PT diagnosis of tumors, viral infection and cellular degeneration and for  
 PT drug screening.  
 XX  
 XX Claim 1; Page 360; 623pp; French.

XX The present invention relates to novel human nucleic acid sequences (I).  
 CC The present sequence is one such nucleic acid sequence. Expression of (I)  
 CC are implicated in tumour suppression or reversion and apoptosis and viral  
 CC resistance. (I) are useful as probes or primers for detecting,  
 CC identifying, measuring and/or amplifying nucleic acid sequences, as  
 CC antisense reagents and for recombinant production of polypeptides. (I),  
 CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
 CC these vectors and antibodies (Ab) against (II) are all useful for  
 CC treatment/prevention of viral, tumour and cell degeneration diseases  
 CC (especially neurodegeneration, such as Alzheimer's disease and  
 CC schizophrenia). Analysing the expression of (I) is also useful for  
 CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
 CC (I) are used for studying the aetiology of these diseases (also immune  
 CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
 CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
 CC in the specification  
 XX

SQ Sequence 374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;  
 Query Match 15.2%; Score 307.4; DB 6; Length 374;  
 Best Local Similarity 88.2%; Pred. No. 3.2e-72;  
 Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 QY 174 TTTCACAGGAGGCGGAAACTTCTGGACCAAAAGCGAGGACATGGAATGGGAACAC 233  
 Db 1 TTTCACAGGCGGAGACCCACTTGGGACCAAAAGCGAGGACATGGAATGGGAGCGC 60  
 QY 234 AGTGAATGAATGACCAAAAGCAAGTATTCATTTTGGTTCTTTGGTTTAAAAAATGTTT 293  
 Db 61 AGTGAATGACATGACCAAAAGCAAGTATTCATTTTGGTTCTTTGGTTTAAAAAATGTTT 120  
 QY 294 ATTTCAGAGTCTTAACACAAAGAAATATTTCTCTGGTGTATTTAAATTTGGTTTGAACA 353  
 Db 121 GTTTGAGTGTCTCAGCACAAAGACATAGCTCTTAGTATTTACTTGGTTCTGGTGGCAGCA 180  
 QY 354 TGAATGGGAAAAGACCAAGGCTGGCAGTGCCTTAAATTTGGAGGAAGGAGCACTTTAG 413  
 Db 181 TGAATGGGAAAAGGACCAAGGCTGGCAGTGCCTTAAATTTGGAGGAAGGAGCACTTTAG 240  
 QY 414 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTTAAATGTTTACTGGAGCAGATGTTGT 473  
 Db 241 TCAACCTCAAGGAAATGTTGGAGAGGCACTTAAATGTTTACTGGAGTANATGTTTGGN 300  
 QY 474 AACAGCTCTTAATGTGCAACTAACACCACTGAAAGAAATTAACCTAAGAGAAATAGCAGA 533  
 Db 301 GACTGCTGNAATGTCNAACCTAACACCACTGAAAGAAATTAACCTGNNAGAAATAGCAGA 360  
 QY 534 AGACAATGAGTGG 546  
 Db 361 GGACAGTGANNNG 373

RESULT 14  
 ABQ94779  
 ID ABQ94779 standard; DNA; 421 BP.  
 XX  
 XX AC ABQ94779;  
 XX  
 XX 28-OCT-2002 (first entry)  
 DT  
 XX Tumour suppression-related oligonucleotide #430.  
 DE  
 XX Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
 KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
 KW viral infection; cell degeneration disease; neurodegeneration; ds;  
 KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX FR2819824-A1.  
 XX  
 XX 26-JUL-2002.  
 XX  
 XX 23-JAN-2001; 2001FR-00000899.  
 XX  
 XX 23-JAN-2001; 2001FR-00000899.  
 XX  
 XX (MOLE-) MOLECULAR ENGINES LAB SA.  
 XX  
 XX Telerman A, Amson R, Tuijnder M, Susini L;  
 XX  
 XX WPI; 2002-610803/66.  
 XX  
 XX New nucleic acid implicated e.g. in tumor suppression, useful for  
 PT diagnosis of tumors, viral infection and cellular degeneration and for  
 PT drug screening.  
 XX  
 XX Claim 1; Page 141; 623pp; French.  
 XX  
 XX The present invention relates to novel human nucleic acid sequences (I).  
 CC



CC The present sequence is one such nucleic acid sequence. Expression of (I)  
CC are implicated in tumour suppression or reversion and apoptosis and viral  
CC resistance. (I) are useful as probes or primers for detecting,  
CC identifying, measuring and/or amplifying nucleic acid sequences, as  
CC antisense reagents and for recombinant production of polypeptides. (I),  
CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
CC these vectors and antibodies (Ab), against (II) are all useful for  
CC treatment/prevention of viral, tumour and cell degeneration diseases  
CC (especially neurodegeneration, such as Alzheimer's disease and  
CC schizophrenia). Analysing the expression of (I) is also useful for  
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
CC (I) are used for studying the aetiology of these diseases (also immune  
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
CC to 280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
CC in the specification  
XX  
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.28; Score 285.8; DB 6; Length 421;  
Best Local Similarity 87.68; Pred. No. 2.2e-66;  
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAAAGCGAGCATGGAATGGGAAACCA 232  
DB 59 CTTTACACGAGGCGGAGACCACTTGGACCAAAAGCGAGCATGGAATGGGAGCG 118  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGTTTAAAAAATGTT 292  
DB 119 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTTTAAAAAATGTT 178  
QY 293 TATTGAGTCTTAACACAAAGATATATCTCTGTGATGTAAATGTTTGTGCAAC 352  
DB 179 TGTTTGAAGTGCTCAGCACAAAGCAATAGCTCTCTAGTAATGTACTTGGTTCGTGAGC 238  
QY 353 ATGAATGGGAAAGACCAAGGCTGGCCTGCACTGCTAAATTTGGAGGAAAGACTTTA 412  
DB 239 ATGAATGGGAAAGACCAAGGCTGGCCTGCACTGCTAAATTTGGAGGAAAGACTTTA 298  
QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGG 472  
DB 299 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGTGGTTGG 358  
QY 473 TACAGCTCTAATGTGCACTAACACCACTGAAAGAAATTAACCTAAGAGAAAT 527  
DB 359 TGACTGCTGTAAATGTCTAATAACCACTGAAAGAAATTAACCTAAGAGAAAT 413

## RESULT 15

ABQ94724  
ID ABQ94724 standard; DNA; 423 BP.

AC ABQ94724;

XX 28-OCT-2002 (first entry)

DE Tumour suppression-related oligonucleotide #375.

XX Tumour; cytostatic; antiviral; neuroprotective; neurotropic; neuroleptic;  
KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
KW viral infection; cell degeneration disease; neurodegeneration; ds;  
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS Homo sapiens.

XX FR2819824-A1.

XX 26-JUL-2002.

XX 23-JAN-2001; 2001FR-00000899.

XX 23-JAN-2001; 2001FR-00000899.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX  
PI  
XX  
DR  
XX  
XX  
PT  
PT  
XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ

Telerman A, Amson R, Tuijnder M, Susini L;  
WPI; 2002-610803/66.

New nucleic acid implicated e.g. in tumor suppression, useful for  
diagnosis of tumors, viral infection and cellular degeneration and for  
drug screening.

Claim 1; Page 125-126; 623pp; French.

The present invention relates to novel human nucleic acid sequences (I).  
The present sequence is one such nucleic acid sequence. Expression of (I)  
are implicated in tumour suppression or reversion and apoptosis and viral  
resistance. (I) are useful as probes or primers for detecting,  
identifying, measuring and/or amplifying nucleic acid sequences, as  
antisense reagents and for recombinant production of polypeptides. (I),  
polypeptides (II) encoded by (I), vector containing (I), cells containing  
these vectors and antibodies (Ab) against (II) are all useful for  
treatment/prevention of viral, tumour and cell degeneration diseases  
(especially neurodegeneration, such as Alzheimer's disease and  
schizophrenia). Analysing the expression of (I) is also useful for  
diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
(I) are used for studying the aetiology of these diseases (also immune  
and inflammatory diseases). Note: In the present specification, SEQ ID 1  
to 280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
in the specification

Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.18; Score 284.8; DB 6; Length 423;  
Best Local Similarity 89.28; Pred. No. 4.2e-66;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAAAGCGAGCATGGAATGGGAAACCA 232  
DB 71 CTTTACACGAGGCGGAGACCACTTGGACCAAAAGCGAGCATGGAATGGGAGCG 130  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGTTTAAAAAATGTT 292  
DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTTTAAAAAATGTT 190  
QY 293 TATTGAGTCTTAACACAAAGATATATCTCTGTGATGTAAATGTTTGTGCAAC 352  
DB 191 TGTTTGAAGTGCTCAGCACAAAGCAATAGCTCTCTAGTAATGTACTTGGTTCGTGAGC 250  
QY 353 ATGAATGGGAAAGACCAAGGCTGGCCTGCACTGCTAAATTTGGAGGAAAGACTTTA 412  
DB 251 ATGAATGGGAAAGACCAAGGCTGGCCTGCTATGCTGATTTGGAGGCAAGACTTTA 310  
QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGG 472  
DB 311 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGTGGTTGG 370  
QY 473 TACAGCTCTAATGTGCACTAACACCACTGAAAGAAATTAACCTAAGAGAAAT 516  
DB 371 TGACTGCTGTAAATGTCTAATAACCACTGAAAGAAATTAACCTAAGAGAAAT 414

Search completed: January 22, 2005, 11:49:54

Job time : 956.5 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds  
(without alignments)  
8723.907 Million cell updates/sec

Title: US-10-069-056-12

Perfect score: 2019

Sequence: 1 atggctggaatgcttactc.....agccgtgactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	107.8	5.3	4680	1	US-08-254-358-1
6	107.8	5.3	4680	1	US-08-475-391-1
7	107.8	5.3	4680	2	US-08-709-609-1
8	107.8	5.3	4680	5	PCT-US95-07178-1
9	107.8	5.3	4681	4	US-09-807-802A-18
10	107.8	5.3	4910	2	US-08-331-384-2
11	107.8	5.3	4910	2	US-08-836-087-2
12	107.8	5.3	4910	3	US-09-246-320-2
13	107.8	5.3	4910	3	US-09-546-738-2
14	107.8	5.3	7214	4	US-09-438-268-1
15	107.8	5.3	7557	4	US-09-770-315-3
16	107.8	5.3	8151	4	US-09-438-268-2
17	107.8	5.3	8179	4	US-09-438-268-5
18	107.8	5.3	8698	4	US-09-770-315-2
19	103	5.1	9339	4	US-09-532-594B-12
20	103	5.1	1197	4	US-09-532-594B-13
21	103	5.1	1611	4	US-09-532-594B-14
22	103	5.1	1872	4	US-09-532-594B-3
23	103	5.1	1872	4	US-09-532-594B-15
24	103	5.1	4767	4	US-09-532-594B-1
25	98.2	4.9	969	4	US-09-807-802A-10
26	98.2	4.9	1200	4	US-09-807-802A-8
27	98.2	4.9	1641	4	US-09-807-802A-6

28	98.2	4.9	1872	4	US-09-807-802A-4	Sequence 4, Appli
29	98.2	4.9	4683	4	US-09-807-802A-19	Sequence 19, Appli
30	98.2	4.9	4718	4	US-09-807-802A-1	Sequence 1, Appli
31	64.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appli
32	60	3.0	1141	4	US-09-806-708B-22	Sequence 22, Appli
33	45.2	2.2	1141	4	US-09-806-708B-22	Sequence 22, Appli
34	40.8	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
35	40.8	2.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
36	40.2	2.0	277	3	US-09-007-005-3	Sequence 3, Appli
37	40.2	2.0	277	3	US-09-244-796-3	Sequence 3, Appli
38	39.4	2.0	580073	4	US-08-545-528B-1	Sequence 1, Appli
39	38.6	1.9	116592	4	US-09-818-512-3	Sequence 3, Appli
40	38.4	1.9	832	4	US-09-621-976-2813	Sequence 2813, Ap
41	38.4	1.9	1367	4	US-09-270-767-14265	Sequence 14265, A
42	38.4	1.9	6328	3	US-08-913-832A-1	Sequence 1, Appli
43	38.4	1.9	6328	4	US-09-249-181A-1	Sequence 1, Appli
44	38.4	1.9	6328	4	US-09-158-707-1	Sequence 1, Appli
45	38.4	1.9	6475	4	US-09-620-312D-325	Sequence 325, App

ALIGNMENTS

RESULT 1  
US-08-336-345-1  
; Sequence 1, Application US/08336345  
; Patent No. 5814510  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,345  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30753  
; REFERENCE/DOCKET NUMBER: 7937-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Parvovirus  
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 1.9e-292;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGTTAAAGGAAAA 60



```

CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,345
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Jennifer
REGISTRATION NUMBER: 30753
REFERENCE/DOCKET NUMBER: 7937-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5049 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Parvovirus
US-08-336-345-2

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.9e-292;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTGTGGGAGCAACCAACTGGTTAAAGGAAAA 60
DB 273 ATGCTGGCACCACTAGTACTGAGGAAGTATGGAGGAGTAAATGGTTAAAGAAACAT 332
QY 61 AGTAACCAAGGAAGTGTCTCAATTTGTTTTAAAAATGAAATGTTCAACTGAATGGA 120
DB 333 GCAGAAAATGAAGCATTTTCGTTTGTGTTTAAATGTGACACGCTCAACTAAATGGAAG 392
QY 121 GATATCGATCGAATAGTTTACAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACA 180
DB 393 GATGTTCTCGCTGGAACTATATACCAACCAATTCACAAATGAAGAGCTAACATCTTTAAT 452
QY 181 CGAGAGCGGAAACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAACCA 234
DB 453 AGAGAGCACAAACAGCAATGGATCAACCGNAGAGAGAAATGCACTGGGATCGGA 512
QY 235 GTGGATGAAATGACCAAAAAGCAAGTATTCATTTTGGATCTTTGGTTAAAAAATGTTTA 294
DB 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGGATGCAATTAATTAAGAAATGTCTT 572
QY 295 TTTGAAGTCTTAAACAAAGATATATTTCTCGTGATGTTAATTTGGTTTGTGCAACAT 354
DB 573 TTTGAAGTCTTTGTTTCTTAAAAATATAGAAACCAATGAATGTGTTTGGTTTATTCAACAT 632
QY 355 GAATGGGAAAGAACCAAGGCTGGCAGTCCCATGTAATTTGGAGGAAAGACATTTAGT 414
DB 633 GAATGGGAAAGATCAAGGCTGGCATTTGTCAATGTTTACTTTCATAGTAGAACTTACA 492
QY 415 CAAGCTCAAGGAAATGTGGAGAGGCAACTAAATATGTTTACTCGAGCAGATGGTTGTA 474
DB 693 CAAGCACTGTTAATGCTACGCAGACAAATGAATATGTAATTTGGATAGATGGTTGGTG 752
QY 475 ACAGCCTGTAATGTGCAACTAACACCAAGCTGAAGAAATTAACCTAAGAGAAATAGACAA 534

```









```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match      5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Qy 1168 GGCAGAAAGAAATGCTGTTTATTTTCATGGACCGACGACAGGCAATCTATTATTGCA 1227
Db      |||||
Qy 1293 GGCAGAGGAAACACCATCTGGCTGTTTGGCGCTGCAACTACCGGGAAGACCAATCGCG 1352
Db      |||||
Qy 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTTATATGCAGCCAAATGTAACCTTT 1287
Db      |||||
Qy 1353 GAGGCCATAGCCACACATGTCGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
Db      |||||
Qy 1288 CCATTTAATGACTGTACCAACAGAACTTGCATTTGGGTAGAGAGCTGTTAACTTTGGA 1347
Db      |||||
Qy 1413 CCCTTCAACGACTGTGTCGACAGATGGTGATCTGGTGGGAGGAGGAAAGATGACCGCC 1472
Db      |||||
Qy 1348 CAGCAAGTAAACACGAGTTTAAAGCCATTGCTCTGGTCAAACTATTTCGCAATGATCAAAA 1407
Db      |||||
Qy 1473 AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGGAAGCAGAGTGGCGGTGGACCAAGAA 1532
Db      |||||
Qy 1408 GGAAGGCGACGAAACAGATTGAACCAACACCGATCATATGACCAAAATGAGAACATT 1467
Db      |||||
Qy 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATGTCACCTCCAACACCAATG 1592
Db      |||||
Qy 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCAACACACTCAACCAATCAGAGACAGA 1527
Db      |||||
Qy 1593 TGGCGCGTGATTGACGGGAACCTCAACGACCTTCGAAACACGAGCGGTTGCAAGACCGG 1652
Db      |||||
Qy 1528 ATGCTTAAACATTCATCTAAACACATACCTTCCTGGTGACTTTGTTTGGTTGACAAAAT 1587
Db      |||||
Qy 1653 ATGTTCAAAATTGAACTCACCCCGCTCTGGATCATGACTTTGGGAAGGTGACCAAGCAG 1712
Db      |||||

Qy 1588 GAA 1590
Db      |||
Qy 1713 GAA 1715
Db      |||
```

## RESULT 6

```
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.
```

```
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-475-391-1
```

```
Query Match      5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Qy 1168 GGCAGAAAGAAATGCTGTTTATTTTCATGGACCGACGACAGGCAATCTATTATTGCA 1227
Db      |||||
Qy 1293 GGCAGAGGAAACACCATCTGGCTGTTTGGCGCTGCAACTACCGGGAAGACCAATCGCG 1352
Db      |||||
Qy 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTTATATGCAGCCAAATGTAACCTTT 1287
Db      |||||
Qy 1353 GAGGCCATAGCCACACATGTCGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
Db      |||||
Qy 1288 CCATTTAATGACTGTACCAACAGAACTTGCATTTGGGTAGAGAGCTGTTAACTTTGGA 1347
Db      |||||
Qy 1413 CCCTTCAACGACTGTGTCGACAGATGGTGATCTGGTGGGAGGAGGAAAGATGACCGCC 1472
Db      |||||
Qy 1348 CAGCAAGTAAACACGAGTTTAAAGCCATTGCTCTGGTCAAACTATTTCGCAATGATCAAAA 1407
Db      |||||
Qy 1473 AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGGAAGCAGAGTGGCGGTGGACCAAGAA 1532
Db      |||||
Qy 1408 GGAAGGCGACGAAACAGATTGAACCAACACCGATCATATGACCAAAATGAGAACATT 1467
Db      |||||
Qy 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATGTCACCTCCAACACCAATG 1592
Db      |||||
Qy 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCAACACACTCAACCAATCAGAGACAGA 1527
Db      |||||
Qy 1593 TGGCGCGTGATTGACGGGAACCTCAACGACCTTCGAAACACGAGCGGTTGCAAGACCGG 1652
Db      |||||
Qy 1528 ATGCTTAAACATTCATCTAAACACATACCTTCCTGGTGACTTTGTTTGGTTGACAAAAT 1587
Db      |||||
Qy 1653 ATGTTCAAAATTGAACTCACCCCGCTCTGGATCATGACTTTGGGAAGGTGACCAAGCAG 1712
Db      |||||

Qy 1588 GAA 1590
Db      |||
Qy 1713 GAA 1715
Db      |||
```

## RESULT 7

```
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```









Db 3539 ATGTTCAAAATTGAACACCCCGCTCTGGATCATGACTTTGGGAAGGTACCAAGCAG 3480  
QY 1588 GAA 1590  
Db 3479 GAA 3477

RESULT 15  
US-09-770-315-3  
; Sequence 3, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 7557  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-3

Query Match 5.3%; Score 107.8; DB 4; Length 7557;  
Best Local Similarity 53.4%; Pred. No. 1.2e-20;  
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 1168 GGCARAAGAAATGCTGTTTATTTCATGACCCAGCCAGCAGCAAAATCTATTATTGCA 1227  
Db 1321 GGCAAGAGGAACACCATCTGGCTGTTTGGCCCTGCAACTACCGGAGACCAACATCGCG 1380  
QY 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTTATTAATGACGCCAATGTAAACTTT 1287  
Db 1381 GAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1440  
QY 1288 CCATTTAATGACTGTACCAACAGAACTTGATTTGGGTAGAAGAGCTGTTAATTGGA 1347  
Db 1441 CCTTCAACGACTGTGTCGACAGATGATGATCTGTGGGAGGAGGGGAAAGATGACCGCC 1500  
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGATTGATCAAAA 1407  
Db 1501 AAGTCGTGGAGTCGGCCAAAGCCATTCTCGAGGAAGCAGGTGCGCGTGGACCAAAA 1560  
QY 1408 GGAAGGCGCAGCAACAGATTGAAACCAACACCACTCATATGACCAAAATGAGAACATT 1467  
Db 1561 TGCAAGTCCTCGGCCAGATAGACCCGACTCCCGTGATGTCACCTCCAAACACCAACATG 1620  
QY 1468 ACAGTGTCTAGATAGCTCGGAAGAAAGACCAACCACTCACTCAACCAATCAGAGACAGA 1527  
Db 1621 TGCCTCGTGTATGACGGGAATCAACGACCTTCGAAACACAGCAGCGGTTGCAAGACCGG 1680  
QY 1528 ATGCTTAACATTCTATCAACATACCTTGCTGCTGCTGCTTTGGTTTGGTTGACAAAAAT 1587  
Db 1681 ATGTTCAAAATTGAACTACCCGCCGCTCTGATCATGACTTTTGGGAAGGTACCAAGCAG 1740  
QY 1588 GAA 1590  
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:05:56  
Job time : 171.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds  
(without alignments)  
10821.780 Million cell updates/sec

Title: US-10-069-056-12  
Perfect score: 2019  
Sequence: 1 agcggcggaatgcttactc.....agccgtgactgactaa 2019

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
  - 2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
  - 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
  - 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
  - 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
  - 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
  - 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
  - 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
  - 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
  - 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
  - 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
  - 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
  - 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
  - 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
  - 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
  - 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
  - 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
  - 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	US-10-647-111A-1	Sequence 1, Appli
2	941.4	46.6	5075	US-10-361-002-3	Sequence 3, Appli
3	941.4	46.6	5075	US-10-361-004-3	Sequence 3, Appli
4	307.4	15.2	374	US-10-466-894-1276	Sequence 1276, Ap
5	285.8	14.2	421	US-10-466-894-430	Sequence 430, App
6	284.8	14.1	423	US-10-466-894-375	Sequence 375, App
7	284.8	14.1	464	US-10-466-894-370	Sequence 370, App
8	284.8	14.1	473	US-10-466-894-367	Sequence 367, App
9	284.8	14.1	486	US-10-466-894-358	Sequence 358, App
10	284.8	14.1	491	US-10-466-894-428	Sequence 428, App
11	284.2	14.1	420	US-10-466-894-438	Sequence 438, App
12	283.8	14.1	343	US-10-466-894-388	Sequence 388, App

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, Ap
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	117	5.8	1386	9	US-09-792-630-32	Sequence 32, Appl
40	117	5.8	1386	10	US-09-953-351-32	Sequence 32, Appl
41	117	5.8	1386	13	US-10-080-376-32	Sequence 32, Appl
42	117	5.8	1386	14	US-10-082-671-38	Sequence 38, Appl
43	117	5.8	1386	14	US-10-097-100-32	Sequence 32, Appl
44	117	5.8	1386	15	US-10-023-208-32	Sequence 32, Appl
45	115.4	5.7	1884	9	US-09-792-630-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1  
US-10-647-111A-1  
; Sequence 1, Application US/10647111A  
; Publication No. US20040209240A1  
; GENERAL INFORMATION:  
; APPLICANT: MALERBA, MADDALENA  
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES  
; FILE REFERENCE: 604-691  
; CURRENT APPLICATION NUMBER: US/10/647,111A  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 5121  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1  
; OTHER INFORMATION: with promoter P4 and left hairpin from WMV (phh1)  
US-10-647-111A-1

Query Match	87.7%	Score	1771	DB	18	Length	5121
Best Local Similarity	92.3%	Pred. No.	0				
Matches	1864	Conservative	0	Mismatches	155	Indels	0
Gaps	0						
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
DB	282	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
QY	61	AGTAACCAAGGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120				
DB	342	AGTAACCAAGGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	401				
QY	121	GATATCGGATGGAATAGTTTACAAAAAAGAGCTGCAGGAGGAGCTGAAATCTTTTACAA	180				

Db 402 GATATCGGATGGATAGTTACAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA 461  
Qy 181 CGAGGAGCGGAAACTACTCTTGGGACCAAAGCAGGACATGAAATGGGAAACCACTGGAT 240  
Db 462 CGAGGAGCGGAAACTACTCTTGGGACCAAAGCAGGACATGAAATGGGAAACCACTGGAT 521  
Qy 241 GAAATGACCAAAAGCAAGCTATTCATTTTGGTTTCTTGGTTTAAATAATGTTTATTGAA 300  
Db 522 GAAATGACCAAAAGCAAGCTATTCATTTTGGTTTCTTGGTTTAAATAATGTTTATTGAA 581  
Qy 301 GTGCTTTAAACAAAGATATATTTCTGCTGATGTTTAAATGTTGTCACCAATGAATGG 360  
Db 582 GTGCTTAAACAAAGATATATTTCTGCTGATGTTTAAATGTTGTCACCAATGAATGG 641  
Qy 361 GGAAGAGCAAGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGACATTTAGTCAAGCT 420  
Db 642 GGAAGAGCAAGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGACATTTAGTCAAGCT 701  
Qy 421 CRAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480  
Db 702 CRAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 761  
Qy 481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGCAAT 540  
Db 762 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGCAAT 821  
Qy 541 GAGTGGGTACTCTACTTACTTATAGCATATAGCAACCAAAAGACATATACCAGTGT 600  
Db 822 GAGTGGGTACTCTACTTACTTATAGCATATAGCAACCAAAAGACATATACCAGTGT 881  
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT 660  
Db 882 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT 941  
Qy 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720  
Db 942 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 1001  
Qy 721 AAAGAGGCGAGCGCCNCTAGTGAGCAACATATACATGATGACATCGGGCCAGAAACG 780  
Db 1002 AAAGAGGCGAGCGCCNCTAGTGAGCAACATATACATGATGATGAAACAGAAACG 1061  
Qy 781 GTTGAACACACAGTAAACCTCGCAGGAAACTAAGCGCGCAGAAATTCAAACTAAAAAA 840  
Db 1062 GTCGAGACCACTGACCTGACAGGAGCTAAGCGCGCAGAAATTCAAACTAGAGAG 1121  
Qy 841 GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGATATAAAGAGTAACCTCACCAGAG 900  
Db 1122 GAGGTCTCGATTAAACCACTCAAAGAGTTGGTACATATAAAGAGTAACCTCACCAGAA 1181  
Qy 901 GACTGGATGATGACGACGACAGTATCAATTGAAATGATGCTCAACCCAGGTGGAGAA 960  
Db 1182 GACTGGATGATGACGACGACAGTATCAATTGAAATGATGCTCAACCCAGGTGGAGAA 1241  
Qy 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAACTTAGCCAGAACCAAAACAGCA 1020  
Db 1242 AACTGCTTAAAAATACACTAGAGATCTGTACACTGACTTAGCAGAAACCAAAACAGCC 1301  
Qy 1021 TTTGACTTAAATTTTAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC 1080  
Db 1302 TTTGACTTAAATTTTAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC 1361  
Qy 1081 ACAAGACCTGCAGAAATTTTGGTTTCTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1140  
Db 1362 ACCAGAACCTGTAGAAATCTTGTGCTAGCAGCTGGCTGGAACTATATAAAGTCTGCCATGCC 1421  
Qy 1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATGCTGTGTTTATTTATGAGCA 1200  
Db 1422 ATCTGTTGTGCTGATAGCAGAGGAGGCAAAAGAACTGTGCTCTTTTACGGACCA 1481  
Qy 1201 GCCAGCACAGCAAACTTATTTATGCAAGGCCATAGCACAGCAGTTGGCAATGTTGTT 1260

Db 1482 GCCAGCACAGGCAAACTATTATTGCACAAGCCATAGCAAGCAGTGTGTTAATGTTGGT 1541  
Qy 1261 TGCTATAATGACGCAAACTGTAAACCTTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320  
Db 1542 TGTTCATATGCTGCCAATGTGNACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1601  
Qy 1321 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCCAGTTTTAAAGCCATTTGCTCT 1380  
Db 1602 TGGGTAGAAAGAGCTGGTAACTTTTGGCCAGCAAGTAAACCCAAATTTCAAAAGCTATTGTTCT 1661  
Qy 1381 GGTCAAACTATTTCGCAATGATCAAAAGAAAGGAGCAGCAACAGATTGAACCAACACCA 1440  
Db 1662 GGCCAAACCATACGCAATGATCAAAAGAAAGGAGCAGCAACAGATTGAACCAACACCA 1721  
Qy 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGGGAGAAAGACCA 1500  
Db 1722 GTTATTATGACCAACCAAGAGNACATTAACCGTGGTTAGNATAGGCTGTGAGGAAAGACCA 1781  
Qy 1501 GAACACACTCAAACCAATCAGAGACAGAAATGCTTTAACTCATCTAAACATACCTTGCTCT 1560  
Db 1782 GAACACACTCAAACCAATCAGAGACAGAAATGCTTTCAACATTCACCTGACACGTAACCT 1841  
Qy 1561 GGTGACTTTGGTTGGTTGACAAATAATGAATGGCCCATGATTTGTGTTGGTTGTAAG 1620  
Db 1842 GGTGACTTTGGTTGGTTGATTAAGCAGCAATGGCCTCTGATCTGCTGCTGGTTGGTGAAG 1901  
Qy 1621 AATGGTTTCAAACTTACCATGGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTTCTGATTGG 1680  
Db 1902 AATGGTTTCAAACTTACCATGGCTTTGTTTACTGTGCTTAAATGGGGCAAAAGTTTCTGATTGG 1961  
Qy 1681 TCAGAAAACTGGGCGAGCCAAAAGTGGCCAACTCTTATATAATTTACTAGTTTGGGACCG 1740  
Db 1962 TCAGAGGACTGGGCGAGCCGAAAGCTAGACACTCTCTATAAAATTCGCTAGGTTCAATGGCG 2021  
Qy 1741 TCACCACTTCACACACCGAAAGTAGCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 2022 TCACCACTCTGACTCCGAGAAAGTAGCCTCTCAGCCAAACTAGCGCTTACTCCACTT 2081  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCCTTGAGAGCGTGTCTTCGGTGGGACCG 1860  
Db 2082 GCATCGGACCTTTGGGACCTAGCTCTAGAGCCTTTGGAGCACCACCAAAATCTCTCTGTTGG 2141  
Qy 1861 GGCATGCGAGAAACCCAGAACTGCGGAGAGCTGTTCCAAAGCCTGCCAAGTGTGCA 1920  
Db 2142 GGCATGCGAGCAAGCCAAACCACTGGGAGGCTGTTCCACAGCCTGCCAAGGTGCTCA 2201  
Qy 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGACCG 1980  
Db 2202 CGAGCCCAACTTTGGTCAGAGATCGAGGCGGATTTGAGAGCTTGTCTTCAGTCAAGAACAG 2261  
Qy 1981 TTGAAGAAAGACTTTCAGGAGCGCGCTGAACCTTGGACTAA 2019  
Db 2262 TTGGAGCGGACTTCAACGAGGAGCTGACCTTGGACTAA 2300

RESULT 2  
US-10-361-002-3  
; Sequence 3, Application US/10361002  
; Publication No. US20040170954A1  
; GENERAL INFORMATION:  
; APPLICANT: Clearant, Inc.  
; APPLICANT: McKenney, Keith  
; APPLICANT: Gillmeister, Lidja  
; APPLICANT: Marlowe, Kristina  
; APPLICANT: Armistead, David  
; TITLE OF INVENTION: Pathogen Inactivation Assay  
; FILE REFERENCE: CI-0043  
; CURRENT APPLICATION NUMBER: US/10/361, 002  
; CURRENT FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 5075



TYPE: DNA  
ORGANISM: Porcine parvovirus  
US-10-361-002-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;  
Best Local Similarity 70.4%; Pred. No. 2.8e-237;  
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

QY	4	GCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAAGT	63
DB	298	GCAGGAAACACTTACTCGGAAGAGGTACTTAAAGCTACCACTGGCTTCAAGATAATGCT	357
QY	64	AACCGAGAGTGTCTCATTTGTTTAAATGAAATGTTCAACTGAATGAAAGAT	123
DB	358	CAAAAAGAGCATTTCTTATGATTTTAAACCAAAAGTCAATCTAAATGGAAAAGAA	417
QY	124	ATCGGATGGAATAGTTTACAAAAAGAGCTGACGAGGACGAGCTGAAATCTTTTACAACGA	183
DB	418	ATTGCTTGGATTAACCTACAACAAGATACAAGATGCGGAATGATAAACCCTACAAGA	477
QY	184	GGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGATGAA	243
DB	478	GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAATCGACAGC	537
QY	244	ATGACCAAAAAGCAAGTATTCATTTTGTGTTTGAATCTTTGGTTAAATAATGTTTATTTGAAGTG	303
DB	538	CTCAGAAAACGGCAAGTACTGATTTTGTGCTCTCTTTGTTTAAATAATGCTCTCTTTGAAGGT	597
QY	304	CTTTACACAAAGATATATTTCTGCTGATGTTAAATTTGTTCTGCAACATGAATGGGA	363
DB	598	ATATTGCAAAAGAACCTAAGTCAAGTGACTGCTGTTCTATCAGCATGAACATGGT	657
QY	364	AAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTGCAAGCTCAA	423
DB	658	CAAGATACTGGCTATCACTGCCATGTACTAGTGGTGAAGAGCTTTACAACAAGCAATG	717
QY	424	GGAAATGGTGAGAGGCACTAAATGTTTACTTGAGCAGATGGTTGGTAAACAGCTGT	483
DB	718	GGAAATGGTTTCAAGAAACAAATTTAAACAATTTTGGAGTAGATGTTTAAATGCAATGC	777
QY	484	AATGTCAACTACACAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAAAGCAATGAG	543
DB	778	AAAGTACTCTAACACAGTTGAAGAAATAAATTAAGGGAATTTAGCAGGATGGTGAG	837
QY	544	TGGTTTACTCTACTTACTTAAAGCATTAAGCAACCAAAAGAGCTATATCAAGTGTGT	603
DB	838	TGGTATCGCTACTAACCCTACACTCACAACAACCTAAAAACAATATACAAAATGACT	897
QY	604	CTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAGCACTAGTCCA	663
DB	898	CATTTTGGAAATATGATGCTTACTACTTCTTAAATAAAAAAGAGCAACT-----	951
QY	664	CCAGAGACGGAGCTATTTCTTACAGTACTCTGGCTGGAAACTTACTTTTAAAA	723
DB	952	GAAAGAGAGCATGGATATATCTCAGCTCAGATTTCTGGCTTCATGACAAATTTCTTAAAA	1011
QY	724	GAAAGGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGAGAAACGGTT	783
DB	1012	GAAAGGAGAGACACTTAGTCACTCAGCTCAGTATTTACTGAAGCAATAAACCTGAACTGTG	1071
QY	784	GAAACCAAGTAAACCTGCGCAGGAACTAAGCGGCGAGAAATCAAACTAAAAAAGAA	843
DB	1072	GAAACCAAGCTTACTCAGCTCAGGAAGCCAAAAGAGGCGAGAAATCAAAACAAAAAAGAA	1131
QY	844	GTTTCTTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACCTCACCAGAGGAC	903
DB	1132	GTAAGCATAAATGCAATTAAGAGCTGGTTTAAATAAAGATGTTACTAGCATAGAAGAC	1191
QY	904	TGATGATGATGACGCGCAGACAGTTTACATTTGAATGATGCTCAACAGGCTGGAGAAAC	963
DB	1192	TGATGATGACAGATCAGACAGTTTATAGAATGATGCTCAACCGGAGAGGAAT	1251
QY	964	CTGCTGAAAAATACGCTAGAGATTTGTGACATACTTACCCAGAACCAAAACAGCATTT	1023

DB	1252	TTAATCAAAAAATACACTAGAAATAACAACCTCTTACTCTAGCAAGAAACAAAAACAGCATAT	1311
QY	1024	GACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTCGCCTGCACA	1083
DB	1312	GACTTAATACTTTGAAAAAGGCAAAACCAAGCATGCTACCAACATTTTAATATTAGCAATACA	1371
QY	1084	AGAACCTGCAGAAATTTTCTGCTTTTATGCTGGAACATATGTTAAAGTTTGCCTGCTATT	1143
DB	1372	AGAACATGTAAATATTTTCAAGCATGCAATTTGGAACATACATTTAAAGCTCTGCCATGTATA	1431
QY	1144	TGCTGTGTTTAAACAGACAGAGGCAAAAGAAATGCTGTTTATTTTTCATGAGCAGGCC	1203
DB	1432	ACTTGTGTTACTAAACAGACAGAGGAAAGAAATACAAATCTTATTTTATGAGGCCAGCA	1491
QY	1204	AGCACAGGCAAAATCTATTATTGCAAGCCATAGCAACAAGCAGTGTGGCAATGTTGTTGC	1263
DB	1492	TCAACAGGAAAGATATAATTGCTCAACACATTTGCAAACTTAGTTGTTGTTGTTGC	1551
QY	1264	TATAATGCAAGCAATGTAAACTTTTCCATTTTAACTGATGCTACCAAGAACTTTGATTGG	1323
DB	1552	TACAAATGCAAGCAATGTAACTTTTCCATTTTAAATGACTGTACAAATTAATAATG	1611
QY	1324	GTAGAGAGCTGCTAACTTTTGGACAGCAAGTAAACAGTAAAGCCATTTTCTCTGCT	1383
DB	1612	ATTGAAGAGGAGGAAACTTTCTTAAACCAAGTAAACCAATTTCAAGGCCATATGTTCAGGT	1671
QY	1384	CAAACTATTTCGCAATTTGATCAAAAGGAAAGGAGGCAAAACAGATTTGAAACCAACAGTC	1443
DB	1672	CAAACTATTGAAATTTGACCAAAAGGTAAGGAGCAAAACAAATTTGAAACCACTCTGTA	1731
QY	1444	ATCATGACCAAAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGCAGAA	1503
DB	1732	ATAATGACTTACAAATGAGACATTAATAAGTTAGAAATAGGATGCGAGGAAAGCAGAA	1791
QY	1504	CACACTCAACCAATCAGACAGCAAGTCTTAACATTTCACTCAACATACCTTTGCTGCTGT	1563
DB	1792	CATACCAACCAATTAAGACAGAAATGTTTAAACATTAACCTAACAGAAACTGCCAGGT	1851
QY	1564	GACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTTGTGCTTGGTTGTTAAAGAAAT	1623
DB	1852	GATTTTGGACTTTTAGAAGAACTGAATGGCCCACTAATATGTTGTTGTTGTTAAAGAA	1911
QY	1624	GTTTACCAATTTACATGCGCAAGCTACTGTCTTAAATGGGGCAAGTTCTCTGATTGCTCA	1683
DB	1912	GGTTACCAAGCAACATGCTAGCTATATGCTCATTTGGGGAATGTACTGATTGCTCA	1971
QY	1684	GAAACTGCGGCGGCAAGGTCCTTCTTAAATTTTACTAGGTTGCGGCAAGCTCA	1743
DB	1972	GAAAAATGGGAGGAGCAAAATGCAAAACCCCAATTAATTAACCA---ACAGACTCTCAG	2028
QY	1744	CCATTCACGACACCGAAAAAGTAGCCCTCTCAGCCAGAACTATGCACTAATCTCCTTGGCA	1803
DB	2028	ATTTCCACATCAGTGAATACTTGGCAGGGGACAACTACGACGCACTCCAAATACAG	2088
QY	1804	TCGGATCTCGA---GGACCTGGCTTTAGAGCTTGGAGCACCACCAAAATCTCC	1853
DB	2089	GAGGACCTGGAATTTAGCTTTAGCTTTGGAGCGGTGGAGGAGGCAACCAACACC	2141

RESULT 3

US-10-361-004-3  
; Sequence 3, Application US/10361004  
; Publication No. US20040170981A1  
; GENERAL INFORMATION:  
; APPLICANT: Clearant, Inc.  
; APPLICANT: McKenney, Keith  
; APPLICANT: Gillmeister, Lidja  
; APPLICANT: Marlowe, Kristina  
; APPLICANT: Armistead, David  
; TITLE OF INVENTION: Real-Time Polymerase Chain Reaction Using Large Target Amplicons  
; FILE REFERENCE: CI-0042  
; CURRENT APPLICATION NUMBER: US/10/361,004

; CURRENT FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 5075  
; TYPE: DNA  
; ORGANISM: Porcine parvovirus  
US-10-361-004-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;  
Best Local Similarity 70.4%; Pred. No. 2.8e-237;  
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

QY 4 GCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAAAGT 63  
DB 298 GCGGGAACACATCTACTCGGAAGAGGTACTAAAGCTACCACTGGCTTCAAGATAATGCT 357

QY 64 AACGAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAAGAAT 123  
DB 358 CAAAAGAGCAATCTCTTATGATTTTAAACACAAAAAGTCAATCTAAATGGAAGAAT 417

QY 124 ATCGGATGGAATAGTTTCAAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTACACGA 183  
DB 418 ATTGCTTGAATTAACACAAAGATACAAAGATGCGGAATGATAAACCTTACAAAGA 477

QY 184 GGAGCGGAACACTTGGGACCAAGCGGAGGACATGGAATGGGAAACCAACAGTGGATGAA 243  
DB 478 GGAGCAGAAACATCATGGGACCAAGGACGACATGGAATGGGAATCAGAAATCGACAGC 537

QY 244 ATGACCAAAAAGCAAGTATTCATTTTGTGTTTGTGTTTAAAAAATGTTTATTTGAAGTG 303  
DB 538 CTCAAAAACGCGAAGTACTGATTTTGTGCTCTTGTAAAAAATGCTCTCTTTGAAGGT 597

QY 304 CTTAACACAAAGAAATATTTCTGGTGATGTTAATTTGTTTGTGCAACATGATGGGA 363  
DB 598 ATATTGCAAAAGAACCTTAAGTCCCAAGTACTGCTGTTTCTATACAGCATGAACATGTT 657

QY 364 AAAGACCAAGCTGGCACTGCCATGTAATTTGGAGGAAGGACATTTAGTCAAGCTCAA 423  
DB 658 CALGATACTGCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717

QY 424 GGAATATGTTGGAGAGGCAATTAATTTTATCTGGAGCAGATGGTTGGTAAACAGCTGT 483  
DB 718 GGAATATGTTTCAGAAACCAATTAACAAATTTATGGAGTAGATGGTTAATTAATGCAATGC 777

QY 484 AATGTGCACTTAACACAGCTGAAGAAATTAACATGAAGAAATAGCAGAGCAATGAG 543  
DB 778 AAAGTACTCTTAACACAGCTGAAGAAATTAACATGAAGAAATAGCAGAGGATGGTGA 837

QY 544 TGGTTTACTCTTACTTATAGCATAGCAACCAACCAACCAACCAACCAACCAACCAACCA 603  
DB 838 TGGTTTACTCTTACTTATAGCATAGCAACCAACCAACCAACCAACCAACCAACCAACCA 897

QY 604 CTTTGGGAACATGATGCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 663  
DB 898 CATTTTGGAAATATGATGCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 951

QY 664 CCAAGAGACGAGCTATTTCTTAGCAGTGACTCTGCTGGTGGAACTACTTCTTAAAAA 723  
DB 952 GAAAGAGAGCATGATGATATTTCTCAGCTCAGATTTCTGCTTCTGCTTCTGCTTCTGCT 1011

QY 724 GAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACCGGTT 783  
DB 1012 GAAGCGAGAGACACTTAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG 1071

QY 784 GAAACCAAGTAACTCTGGCAGGAACTAAGCGCGGAGAAATTCAACTAAAAAGAA 843  
DB 1072 GAAACCAAGTAACTCTGGCAGGAACTAAGCGCGGAGAAATTCAACTAAAAAGAA 1131

QY 844 GTTCTTATTAACACTACACTTAAAGAGCTGGTGATTAAGAGTAACTCACCAGAGGAC 903  
DB 1132 GTAAGCATAAATGCACAAATGAAGAGCTTGGTTTAAATGAAGATGCTAGCATGAAGAC 1191

## RESULT 4

US-10-466-894-1276  
; Sequence 1276, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius

```
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-1276

Query Match      15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 1.8e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGCGGAACTACTTTGGACCAAGCGAGACATGGAATGGAAACCCAC 233
DB 1 TTTTCAACGAGGCGGAGACCACTTTGGACCAAGCGAGACATGGAATGGAGAGCGC 60

QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATCTTTGGTAAATAATGTTT 293
DB 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATCTTTGGTAAAGAGTGT 120

QY 294 ATTGAAGTCTTAAACAAAGATATATTTCTGCTGATGTTTAAATGTTTGTGCAACA 353
DB 121 GTTTGAAGTCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTTCTGTCGAGCA 180

QY 354 TGAATGGGAAAGACCAAGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGCACTTTAG 413
DB 181 TGAATGGGAAAGACCAAGCTGGCAGCTGCTCATGTCTGATTTGGAGGCAAGCACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGTTGGAGAGCAAGCACTTAATTTTACTGAGCAGATGTTGGT 473
DB 241 TCAACCTCAAGGAAATGTTGGAGAGCAAGCACTTAATTTGCTAGTANATGTTGGN 300

QY 474 AACAGCTGTATGTCGAATACACCAAGCTGGAAGATTAAGCAATTAAGCAATAGCAGA 533
DB 301 GACTGCTGNAATGTNCAACTAACCAAGCTGGAAGATTTAAACTGNNAGAAATAGCAGA 360

QY 534 AGACAATGAGTGG 546
DB 361 GGACAGTGANNGG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Teلمان, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1275
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```

```
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 276, 405, 417, 419, 421
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-430

Query Match      14.2%; Score 285.8; DB 18; Length 421;
Best Local Similarity 87.6%; Pred. No. 9.8e-65;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 173 CTTTCAACGAGGCGGAACTACTTTGGACCAAGCGAGACATGGAATGGAAACCA 232
DB 59 CTTTCAACGAGGCGGAGACCACTTTGGACCAAGCGAGACATGGAATGGAGAGCG 118

QY 233 CAGTGATGAAATGACCAAAAGCAAGTATTCATTTTGTGATCTTTGGTAAATAATGTT 292
DB 119 CAGTGATGACATGACCAAAAGCAAGTATTTATTTTGTGATCTTTGGTAAAGTGT 178

QY 293 TATTTGAAGTCTTAAACAAAGATATATTTCTGCTGATGTTAAATGTTTGTGCAAC 352
DB 179 TGTTTGAAGTCTCAGCACAAAGACATAGCTCTCTAGTAATGTTTCTGTCGAGC 238

QY 353 ATGAATGGGAAAGCAAGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGCACTTTA 412
DB 239 ATGAATGGGAAAGCAAGCTGGCAGCTGCTCATGTCTGATTTGGAGGCAAGCACTTTA 298

QY 413 GTCAAGCTCAAGGAAATGTTGGAGAGCAACTAAATGTTTCTGAGCAGATGTTGG 472
DB 299 GTCAACCTCAAGAAATGTTGGAGAGCAAGCACTAAATGTTTCTGAGTAGATGTTGG 358

QY 473 TAACAGCTGTATGTCGAACCTAACCAAGCTGAAAGAAATTAACCTAAGAGAAAT 527
DB 359 TGACTGCTGTATGTTCAACTAACCAAGCTGAAAGAAATTAACCTAAGAGAAAT 413

RESULT 6
US-10-466-894-375
; Sequence 375, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Teلمان, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```

Query Match 14.1%; Score 284.8; DB 18; Length 423;  
Best Local Similarity 89.2%; Pred. No. 1.9e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAACCA 232  
DB 71 CTTTACACGAGGCGGAGACACCTTTGGGACCAAGCGAGGACATGGAATGGGAGCG 130

QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCATTTTTCATTTTGGTTAAAGATGTT 292  
DB 131 CAGTGAATGACATGACCAAAAGCAAGTATTCATTTTTCATTTTGGTTAAAGATGTT 190

QY 293 TATTTCAAGTCTTAACACAAAGAAATATTTCTGCTGATGTTAATTTGGTTTGTGCAAC 352  
DB 191 TGTTGAAGTCTCAGCACAAAGAAATAGTCTTCTGCTGATGTTAATTTGGTTTGTGCAAC 250

QY 353 ATGAATGGGAAAGCAAGGCTGCGACTGCAATGCTAATTTGGGAAAGGACTTTA 412  
DB 251 ATGAATGGGAAAGCAAGGCTGCGACTGCTGCTGATGTTGAGGCAAGGACTTTA 310

QY 413 GTCAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTTACTGGACAGATGTTGG 472  
DB 311 GTCAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTTACTGGAGTAGATGTTGG 370

QY 473 TAACAGCTCTAATGTGCAACTAACACAGCTGAAAGAAATTTAA 516  
DB 371 TGACTGCTGTAAATGTTCAACTAACACAGCTGAAAGAAATTTAA 414

RESULT 7  
US-10-466-894-370  
; Sequence 370, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Teleman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466, 894  
; PRIOR FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 370  
; LENGTH: 464  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 414, 435, 459, 462, 464  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-370

Query Match 14.1%; Score 284.8; DB 18; Length 464;  
Best Local Similarity 89.2%; Pred. No. 1.9e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAACCA 232  
DB 55 CTTTACACGAGGCGGAGACCACTTTGGGACCAAGCGAGGACATGGAATGGGAGCG 114

QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCATTTTGGTTTGTGTTAAAGATGTT 292  
DB 115 CAGTGAATGACATGACCAAAAGCAAGTATTCATTTTGGTTTGTGTTAAGATGTT 174

QY 293 TATTTGAAGTCTTAACACAAAGAAATATATTTCTGCTGATGTTAATTTGGTTTGTGCAAC 352  
DB 175 TGTTTGAAGTCTCAGCACAAAGAAACATAGCTCCCTAGTAATGTTTACTTTGGTTGTCGAGC 234

QY 353 ATGAATGGGAAAGCAAGGCTGCGACTGCAATGCTAATTTGGGAAAGGACTTTA 412  
DB 235 ATGAATGGGAAAGCAAGGCTGCGACTGCTGCTGATGTTGAGGCAAGGACTTTA 294

QY 413 GTCAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTTACTGGACAGATGTTGG 472  
DB 295 GTCAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTTACTGGAGTAGATGTTGG 354

QY 473 TAACAGCTCTAATGTGCAACTAACACAGCTGAAAGAAATTTAA 516  
DB 355 TGACTGCTGTAAATGTTCAACTAACACAGCTGAAAGAAATTTAA 398

RESULT 8  
US-10-466-894-367  
; Sequence 367, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Teleman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466, 894  
; PRIOR FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 367  
; LENGTH: 473  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 10, 431, 447, 467, 468, 473  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-367

Query Match 14.1%; Score 284.8; DB 18; Length 473;  
Best Local Similarity 89.2%; Pred. No. 1.9e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAACCA 232  
DB 72 CTTTACACGAGGCGGAGACCACTTTGGGACCAAGCGAGGACATGGAATGGGAGCG 131

QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCATTTTTCATTTTGGTTTAAAGATGTT 292  
DB 132 CAGTGAATGACATGACCAAAAGCAAGTATTCATTTTTCATTTTGGTTTAAAGATGTT 191

QY 293 TATTTGAAGTCTTAACACAAAGAAATATTTCTGCTGATGTTAATTTGGTTTGTGCAAC 352  
DB 192 TGTTTGAAGTCTCAGCACAAAGAAATAGCTCTCTAGTAATGTTTACTTTGGTTGTCGAGC 251

QY 353 ATGAATGGGAAAGCAAGGCTGCGACTGCAATGCTAATTTGGGAAAGGACTTTA 412  
DB 252 ATGAATGGGAAAGCAAGGCTGCGACTGCTGCTGATGTTGAGGCAAGGACTTTA 311

QY 413 GTCAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTTACTGGACAGATGTTGG 472  
DB 312 GTCAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTTACTGGAGTAGATGTTGG 371

QY 473 TAACAGCTCTAATGTGCAACTAACACAGCTGAAAGAAATTTAA 516

Db 372 TGACTGCTGTAAATGTTCAACTAACACGAGTGAACAAAAA 415  
|||||

## RESULT 9

US-10-466-894-358  
; Sequence 358, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 358  
; LENGTH: 486  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 413, 428, 435, 448, 457, 462  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;  
Best Local Similarity 89.2%; Pred. No. 1.9e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
  
QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 232  
|||  
Db 54 CTTTACACGAGGAGCGGAGCACACTTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 113  
|||  
QY 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTAAAAAATGTT 292  
|||  
Db 114 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGGATTTCTTTGGTTAAAGAAGTGT 173  
|||  
QY 293 TATTTGAAGTGCTTAACACAAAGATATTTCTCTGTGTATTTAAATGTTTGTGCAAC 352  
|||  
Db 174 TGTTTGAAGTGCTCAGCACAAAGAACATAGCTCCTAGTAAATGTTTGTGTCGAGC 233  
|||  
QY 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCACTGTAATTTGGAGGAAAGACTTTA 412  
|||  
Db 234 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCATGCTGATTTGGAGGCAAGACTTTA 293  
|||  
QY 413 GTCAAGCTCAAGGAAATGTTGGAGAAAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
|||  
Db 294 GTCAACCTCAAGGAAATGTTGGAGAAAGCAAGTAAATGTTTACTGGAGTAGATGTTGG 353  
|||  
QY 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAGAAATTTAA 516  
|||  
Db 354 TGACTGCTGTAAATGTTCAACTAACACGAGTGAACAAAAA 397  
|||

## RESULT 10

US-10-466-894-428  
; Sequence 428, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 428  
; LENGTH: 491  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 12, 421, 441  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;  
Best Local Similarity 89.2%; Pred. No. 2e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
  
QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 232  
|||  
Db 62 CTTTACACGAGGAGCGGAGCACACTTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 121  
|||  
QY 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTAAAAAATGTT 292  
|||  
Db 122 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGGATTTCTTTGGTTAAAGAAGTGT 181  
|||  
QY 293 TATTTGAAGTGCTTAACACAAAGATATTTCTCTGTGTATTTAAATGTTTGTGCAAC 352  
|||  
Db 182 TGTTTGAAGTGCTCAGCACAAAGAACATAGCTCCTAGTAAATGTTTGTGTCGAGC 241  
|||  
QY 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCACTGTAATTTGGAGGAAAGACTTTA 412  
|||  
Db 242 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCATGCTGATTTGGAGGCAAGACTTTA 301  
|||  
QY 413 GTCAAGCTCAAGGAAATGTTGGAGAAAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472  
|||  
Db 302 GTCAACCTCAAGGAAATGTTGGAGAAAGCAAGTAAATGTTTACTGGAGTAGATGTTGG 361  
|||  
QY 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAGAAATTTAA 516  
|||  
Db 362 TGACTGCTGTAAATGTTCAACTAACACAGCTGAAGAAATTTAA 405  
|||

## RESULT 11

US-10-466-894-438  
; Sequence 438, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 438





***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:58 , Search time 6354.75 Seconds  
(without alignments)  
11577.453 Million cell updates/sec

Title: US-10-069-056-12  
Perfect score: 2019  
Sequence: 1 atggctggaatgcttactc.....agccgtgacttggaactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gssi:\*  
9: gb\_gssi2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	61.6	3.1	1101	9	CNS0039G
C 2	59.8	3.0	997	9	CNS005TE
C 3	50	2.5	1134	8	CC245800
C 4	50	2.5	1241	9	AG448181
C 5	49.8	2.5	1101	9	CNS0006J
C 6	49.8	2.5	1101	9	CNS0182P
C 7	49.6	2.5	1101	9	CNS017XK
C 8	49.4	2.4	886	9	CNS075DF
C 9	49.2	2.4	913	9	CG772152
C 10	49.2	2.4	930	9	CG770435
C 11	48.4	2.4	1083	9	CNS07A2Y
C 12	47.8	2.4	987	9	CNS00418
C 13	47.6	2.4	1025	9	CNS014J2
C 14	47.2	2.3	1811	9	CG753732
C 15	46.6	2.3	933	9	CNS00CFN
C 16	46.4	2.3	1080	9	CNS00EPP
C 17	46.2	2.3	1758	9	CL509408
C 18	46	2.3	769	9	CNS011WN
C 19	45.8	2.3	963	9	CNS00A4L
C 20	45.8	2.3	1066	6	CD386086
C 21	45.8	2.3	1204	9	CNS016E2
C 22	45.6	2.3	1200	9	CNS016EL
C 23	45.4	2.2	975	6	CD050911
C 24	45.2	2.2	644	8	BH463121

BOHEK14TR

Query Match 3.1%; Score 61.6; DB 9; Length 1101;  
Best Local Similarity 16.2%; Pred. No. 0.00018;

C 25	45.2	2.2	979	8	AZ538989
C 26	45	2.2	768	9	CNS01VSE
C 27	45	2.2	894	9	CNS018BG
C 28	45	2.2	1184	8	B13117
C 29	44.8	2.2	368	2	AW767845
C 30	44.8	2.2	414	4	BJ640957
C 31	44.8	2.2	435	8	BZ665713
C 32	44.8	2.2	530	9	CE641746
C 33	44.8	2.2	581	9	CL553259
C 34	44.8	2.2	1843	9	AG435185
C 35	44.6	2.2	298	4	B1815258
C 36	44.6	2.2	363	5	BQ451492
C 37	44.6	2.2	870	8	AZ693037
C 38	44.6	2.2	1199	8	CC240969
C 39	44.4	2.2	706	9	AG178839
C 40	44.4	2.2	878	9	CNS0187R
C 41	44.4	2.2	938	9	CNS006TJ
C 42	44.4	2.2	1201	5	BX416077
C 43	44.2	2.2	588	1	AJ773389
C 44	44.2	2.2	694	1	AJ778527
C 45	44.2	2.2	1500	9	AG334092

#### ALIGNMENTS

RESULT 1  
CNS0039G/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL063921  
VERSION  
AL063921.1 GI:4941778  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 1101)  
Genoscope.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
JOURNAL  
COMMENT  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/TheBDGP/Drosophila> melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoeer in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
Location/Qualifiers  
1..1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR08K10"  
/clone\_lib="RPCI-98"  
/notes="end : TET3"

source  
1..1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR08K10"  
/clone\_lib="RPCI-98"  
/notes="end : TET3"

#### ORIGIN

Query Match 3.1%; Score 61.6; DB 9; Length 1101;  
Best Local Similarity 16.2%; Pred. No. 0.00018;

Matches 102; Conservative 266; Mismatches 260; Indels 0; Gaps 0;

QY 60 AACTACACGAGAGTGTCTCAATTTGTTTAAATGAAATGTTCAACTGAATGAA 119  
 Db 1099 ARKGGDTWDRTKDDWTKWTTWKDRADRRWAGDADRWDGAGTGTATWWW 1040  
 QY 120 AGATATCGGATGAATAGTTTACAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTACA 179  
 Db 1039 WWWWATWDTWDDKWWATAAATDTATWRTAWRADWAGDRGAGKRDRAADADGAG 980  
 QY 180 ACAGAGAGCGGAACACTACTTGGGACCAACCGGAGGACATGGAATGGAAACACAGTGA 239  
 Db 979 RDGGRKRDKDKRDGDDDKGGKKKAKAATKAWDKWDDWDDKWDGAKRKAD 920  
 QY 240 TGAATAGCACCACAAAGCAAGTATTCTTTGTTGTTGTTTAAATAATGTTTATTGA 299  
 Db 919 DDGAGGKDDGKGDADDDTGTGDDDKDDWDDKAKGTGWDATWAAATDWWWG 860  
 QY 300 AGTGTCTTAACACAAAGATATATTTCTCGTGTGATGTTAAATGTTGTTGCAACATGAAT 359  
 Db 859 ADADWTTWDAADDDWADDDWDAWAKWDDAWAGARTADRDWDGDRAGKRGARKRRDR 800  
 QY 360 GGAAGAGCAGGAGCTGCGACATGCTACTTAATTTGGAGGAGAGCACTTTAGTCAAGC 419  
 Db 799 KRADDKRDADDDAATWTTTDTTDDWKKWTDWTRWAADRTWDRDDDDDRDRA 740  
 QY 420 TCAAGGGAATGTGGAGAGGCAACTAAATGTTTCTGAGCAGATGTTGTTGTAACAGC 479  
 Db 739 GTAGRWRTWKRRKRDTRWDADADDTARDRRRRGGDDGADAGKKGKTKGRRRRDR 680  
 QY 480 CTGTAATGTGCAACTAACACAGCTGAAAGAAATTAACACTAAGAGAAATAGCAGAGACAA 539  
 Db 679 ATWDRTDAAWAAWTTTDTDDWDRKRRKRRRRRTTARAADWWTWKAWDWAK 620  
 QY 540 TGAGTGGTCTACTTACTTACTTATAGCATAGCAACCAACAAAGAGACTATACCAAGTG 599  
 Db 619 WDKTRADRWDAADTWDTRADKADRWAKARAWRRARRDRARAADRRTTKTKTTAT 560  
 QY 600 TGTCTCTTTTGGAAACATGATGCTTACTTACTTATTTTAACTAAAGAAATAAGCACTAG 659  
 Db 559 WTTWAARAANAANAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWA 500  
 QY 660 TCACCAAGAGAGCGGAGCTTATTTCTT 687  
 Db 499 AWTAAWAAAAAATAATTTTTTTTTT 472

## RESULT 2

CNS005TE/c

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR12K22 of RCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 997)  
 Genoscope.

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and

Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source

1..997  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR12K22"  
 /clone\_lib="RCI-98"  
 /note="end : TET3"

## ORIGIN

Query Match 3.0%; Score 59.8; DB 9; Length 997;

Best Local Similarity 19.1%; Pred. No. 0.00051;

Matches 167; Conservative 183; Mismatches 526; Indels 0; Gaps 0;

QY 25 GAAGTTTGGAGCAACCAACTGGTTAAAGGAAAAAGTAACAGGAAGTCTCTCATTT 84  
 Db 963 RRARRRRARGARAGARAGARAGARAGARAGARAGARAGARAGARAGARAGARAG 904  
 QY 85 GTTTTAAATGAAAAATGTTCAATGAAATGGAAGATATCGATGGAATAGTTCAAA 144  
 Db 903 RGGGRRRAAAR 844  
 QY 145 AAAGAGCTCGAGGAGCAGCTGAAATCTTTAAACAGGAGCGGAAACTACTTGGAC 204  
 Db 843 GAGGRAGRGRGRRRRRARRRAGAARARARGRRARARARRRRRARRRARGARG 784  
 QY 205 CAAAGCAGGACATGGAATGGGAAACACAGTGGATGAATGACCAAGCAAGTATTC 264  
 Db 783 RRRGRRGRGRGRRRRRARRRARRARAGARAGAAARRRRARRRARRRRRRRR 724  
 QY 265 ATTTTGTATTCTTTGGTTAAAAATGTTTATTTGAAGTCTTAAACACAAAGATATATT 324  
 Db 723 RRGAGARARRRRARRARARAGAGARRRRRRGRRGAGARRRRGRRRGAGARRRR 664  
 QY 325 CTGTTGATGTTAAATGTTGTTGCAACATGAATGGGAAAGCAAGCTGGACTGC 384  
 Db 663 GRMRTRRRRRRRARRGAAARRRARGARRRRRRRRRRRRRRRRRRRRRRRRRR 604  
 QY 385 CATGTACTAATTGGAGAAAGGACTTTAGTCAAGCTCAAGGGAATGGTGGAGAGGCAA 444  
 Db 603 RARGARRRRRRRRRRRRRRRAGARRRRRRRRRRRRRRRRRRRRRRRRRRRRRA 544  
 QY 445 CTAAATGTTTACTGAGCAGATGTTGGTAAACAGCCTGTAATGTGCAACTAACACCACT 504  
 Db 543 ARGAGAGARAGRAGARRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRAARA 484  
 QY 505 GAAAGAAATAACTAAGAGAAATAGCAGAGACAATGAGTGGGTACTCTTACTTACTTAT 564  
 Db 483 AARARRRARRAGARR 424  
 QY 565 AACATAAGCAACCAAAAGAAAGACTATACCAAGTGTCTTCTTTTGGAAACATGATGCT 624  
 Db 423 NNN 364  
 QY 625 TACTATTTTTAACTAAAAAGAAAAAAGCACTAGTCCACCAAGAGAGCGGAGGCTATTTT 684  
 Db 363 NTTNNNNNTTNT 304  
 QY 685 CTTAGCAGTGAATCTGCTGGTGGAAAACTAACTTTTAAAGAGCGGCGCCCTCTAGTG 744  
 Db 303 MTTNNNNNAAANNATNANAAAAAANANNNNNNNNNNNNNNNNNNNNNNNNNNNNN 244  
 QY 745 AGCAAACTATACACTGATGACATGCGGCCAGAAACGGTTGAAACACCAAGTAAACCTGCG 804

[illegible]



CNS0182P 1101 bp DNA linear GSS 26-JUL-1999  
Drosophila melanogaster genome survey sequence SP6 end of BAC  
BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL108811  
VERSION AL108811.1 GI:5629115  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
source  
1. .1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACN37D10"  
/clone\_lib="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : SP6"

ORIGIN  
Query Match 2.5%; Score 49.8; DB 9; Length 1101;  
Best Local Similarity 21.9%; Pred. No. 0.2;  
Matches 115; Conservative 179; Mismatches 228; Indels 3; Gaps 1;  
207 AACGAGGACATGGATGGAAACACACAGTGGATGAATGACCAAAAGCAAGTATTCAT 266  
217 AAGGAAAAAANAANAANAANAANAAGGKGAAGGGRGRRBVGAGDAAATAAA 576  
227 TTTTGAATCTTGGTTAAAAAATGTTTATTGAAGTCTTAACACAAAGAAATATATTC 326  
237 AAADTKORWDARGAARAARAADAARAARWAARWAARAARAARAARAARAARAAG 636  
247 TGGTGATGTTAAATGTTGTCAACATGAATGGGAAAGACCAAGGCTGGCACTGCCA 386  
257 TTCTAAAAAAGGGGKTKGKAADGDDGAAADAARWRGRGRRAARAARAADAAR 696  
267 TGTACTAATTTGGAGGAAGACATTTAGTCAAGCTCAAGGAAATGTTGGAGAGCACT 446  
277 AAAAAAANAANAANAANAAGGKKGKDGKATKTAARAAKGRDGTATATMTWD 756  
287 AAATGTTTACTGGACGAGATGGTTGTTAAGCTGTAATGTGCAACTCAACACAGCTGA 506  
297 ATWATWKAATDTAKRAAAGRRKRDKARTARGRRARTR---RRAAAGGRRARAGAR 813  
307 AAGAATTAATAAGAGAAATAGCAGAGACAATGAGTGGGTTACTCTACTTACTATTA 566  
317 AAAAAAADAARDWAANAANAANAANAANAANAANAANAANAANAANAANAANA 873  
327 GCATAGCAAAACCAAAAGACTATACCAAGTGTGTTCTTTTGGAAACATGATGCTTA 626  
337 RARARRRRRRRRRARRARAADAATDTKRWADATTDKTTTWTDDDDWDKAKRD 933  
347 CTATTTTAACTAAAGAAGAAATAGCACTAGTCCACCAAGAGACGAGGCTATTTC 686  
357 RWAAXADGAWKWRDRAARDWAATAKDDDGWKWGGGRKRGKDKKRWKGTGKTKDD 993

QY 687 TAGCACTGACTCTGGCTGGAAACTAACTTTTAAAGAGGCGA 731  
Db 994 DWDKTTWRDWWWWTRTKWDWMDGGRGRWRTRKRWGAWWRADA 1038

RESULT 7  
CNS017KX/c  
LOCUS  
DEFINITION  
BACN37F10 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL108171  
VERSION AL108171.1 GI:5628475  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
Location/Qualifiers  
1. .1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACN37F10"  
/clone\_lib="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : SP6"

ORIGIN  
Query Match 2.5%; Score 49.6; DB 9; Length 1101;  
Best Local Similarity 21.8%; Pred. No. 0.23;  
Matches 95; Conservative 143; Mismatches 198; Indels 0; Gaps 0;  
260 TATTCATTTTGTATCTTTGGTTAAAAAATGTTTATTGAGTGTCTTAACACAAAGATA 319  
1077 TWKTADKKWTRDWTDDGGGGGGGKGGKGGKADTKRDMRGATRRARWATAATATA 1018  
320 TATTTCTCTGTGATGTTAATGTTTGTCAACATGAATGGGAAAGACCAAGGCTGGC 379  
1017 DDTWKKTGTGTTKTKWKKWARRAARWDRDWTDDWMDTKWDRWTRWTRATRDGR 958  
380 ACTGCATCTACTAATTTGGAGGAAAGCACTTTAGTCAAGCTCAAGGAAATGTTGGAGAA 439  
957 AGRARERDRDRGRDRRRRRRRTRKWRWRRAAWTATTTWWGTTKGAWTKRKGK 898  
440 GGCAACTAATGTTTACTGAGCAGATGTTGTTGTAACAGCTGTATGTCACACTAACAC 499  
897 GAAAAADWGTDDWRDKRAAAAAAGKGGKGAATWGAAAAAAMWDTTKTKTKRT 838  
500 CAGCTGAAGAAATTAACCTAAGAGAATAGCAGACATGATGGTGTCTACTCTACTTA 559  
837 KTGRAWAAAGTTRTWDAAATWATKRAAAAAAGRAAAAWAKDRGTGKRGKRGKGTDDGK 778  
560 CTTATAAGCAATAGCAAAACCAAAAAAGACTATATCAAGAGTGTGTTCTTTTGGAAAAATGA 619  
777 TWKTTTWTWTAARARARARAAAKGWTKTGDRBARAWDANTTTTRKGAATRDGAA 718  
620 TTGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGTCCACCAAGAGACGGAGCT 679



<div style="text-align: right;">/strain="GA-2"</div> <div style="text-align: right;">/db xref="taxon:7070"</div> <div style="text-align: right;">/clone lib="Tribolium BAC library"</div> <div style="text-align: right;">/note="Vector: pBAC3.6; Site 1: EcorI; Site 2: EcorI;</div> <div style="text-align: right;">Library constructed by Exelixis Inc."</div>					
ORIGIN					
Query Match            2.4%; Score 49.2; DB 9; Length 913;					
Best Local Similarity 58.0%; Pred. No. 0.27;					
Matches     87; Conservative     0; Mismatches     63; Indels     0; Gaps     0;					
QY	23	ATGAGCTTTGGGAGCAACCAACTGGTTTAAGGAAAAAAGTAACCGAGGAACTGTTCAT	82		
DB	296	ATGAATTTTACACACAGACAATAATATACATCTTTAAATTGATTTGGAAGTATCTGAA	355		
QY	83	TTGTTTTTAAAAATGAAAATGTTCCAACTGAATGAAAAAGATATCGGATGGAATAGTTACA	142		
DB	356	CTATTTTATAACAAAAATCTTCGAAGAACAGAAAAATAAACTATCGACTAACTAAA	415		
QY	143	AAAAAGAGCTGCAGGAGCAGCTGAAAT	172		
DB	416	AAATAACTATCGACAACCTCTGTACTAAAAAT	445		
RESULT 10	CG770435            930 bp     DNA     linear     GSS 29-OCT-2003				
LOCUS	TCB52.2 A06.SP6 Tribolium BAC library Tribolium castaneum genomic,				
DEFINITION	genomic survey sequence.				
ACCESSION	CG770435				
VERSION	CG770435.1 GI:38023614				
KEYWORDS	GSS.				
SOURCE	Tribolium castaneum (red flour beetle)				
ORGANISM	Tribolium castaneum Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Tenebrionidae; Tribolium. 1 (bases 1 to 930) Savard,J. and Tautz,D. Tribolium castaneum BAC-ends sequencing project Unpublished (2003) Contact: Savard, J. Abteilung fur Evolutionsgenetik, AG Tautz Institut fur Genetik, Universitat zu Koln Weyertal 121, 50931 Koln, Germany Tel.: 49 221 470 6911 Fax: 49 221 470 5975 Email: savard@uni-koeln.de Class: BAC ends.				
REFERENCE	Location/Qualifiers				
AUTHORS	1. .930				
TITLE	/organism="Tribolium castaneum"				
JOURNAL	/mol_type="genomic DNA"				
COMMENT	/strain="GA-2"				
FEATURES	/db xref="taxon:7070"				
source	/clone lib="Tribolium BAC library"				
	/note="Vector: pBAC3.6; Site 1: EcorI; Site 2: EcorI;				
	Library constructed by Exelixis Inc."				
ORIGIN					
Query Match            2.4%; Score 49.2; DB 9; Length 930;					
Best Local Similarity 58.0%; Pred. No. 0.28;					
Matches     87; Conservative     0; Mismatches     63; Indels     0; Gaps     0;					
QY	23	ATGAGCTTTGGGAGCAACCAACTGGTTTAAGGAAAAAAGTAACCGAGGAACTGTTCAT	82		
DB	288	ATGAATTTTACACACAGACAATAATATACATCTTTAAATTGATTTGGAAGTATCTGAA	347		
QY	83	TTGTTTTTAAAAATGAAAATGTTCCAACTGAATGAAAAAGATATCGGATGGAATAGTTACA	142		
DB	348	CTATTTTATAACAAAAATCTTCGAAGAACAGAAAAATAAACTATCGACTAACTAAA	407		
QY	143	AAAAAGAGCTGCAGGAGCAGCTGAAAT	172		

Db							408	AAATAACTATCGACAACCTGTGACTTAAAT	437
RESULT 11	CNS07A2Y								
LOCUS	CNS07A2Y								
DEFINITION	T3 end of clone BC0A001B01 of library BCOAA from strain CBS 767 of Debaryomyces hansenii, genomic survey sequence.								
ACCESSION	AL436064								
KEYWORDS	AL436064.1 GI:12219477								
SOURCE	GSS								
ORGANISM	Debaryomyces hansenii (anamorph: Candida famata)								
REFERENCE	Debaryomyces hansenii								
AUTHORS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Debaryomyces. 1 (bases 1 to 1063) Souciet-J.L., Aigle,M., Artiguenave,P., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrien,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Olier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Winker,P. and Weissenbach,J. Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies FEMS Lett. 487 (1), 3-12 (2000)								
TITLE	yeast species for molecular evolution studies								
JOURNAL	FEMS Lett. 487 (1), 3-12 (2000)								
MEDLINE	20584711								
PUBMED	11152876								
REFERENCE	2 (bases 1 to 1063)								
AUTHORS	Lepingle,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H., Artiguenave,F., Wincker,P. and Gaillardin,C. Genomic exploration of the hemiascomycetous yeasts: 14. Debaryomyces hansenii var. hansenii FEMS Lett. 487 (1), 82-85 (2000)								
JOURNAL	20584724								
MEDLINE	11152889								
PUBMED	11152889								
REFERENCE	3 (bases 1 to 1063)								
AUTHORS	Genoscope.								
TITLE	Direct Submission								
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)								
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.								
FEATURES	Location/Qualifiers								
source	1..1063 /organism="Debaryomyces hansenii" /mol_type="genomic DNA" /strain="CBS 767" /variety="hansenii" /db_xref="taxon:4959" /clone="BC0A001B01" /clone_lib="BCOAA" /note="end : T3"								
ORIGIN									
Query Match	2.4%; Score 48.4; DB 9; Length 1063;								
Best Local Similarity	38.2%; Pred.No. 0.46; Indels 0; Gaps 0;								
Matches 108; Conservative	38; Mismatches 137;								
QY	369 CCAAGCCTGGCATGTCATTAATTGGAGGAAGGACTTTAGTCAAGCTCAAGGAA 428								
Db	661 CCCGCSSSSACC GG GTTG TTTCTGTGSGSTTTTTTTTTWWWKWWATWWTAT 720								
QY	429 ATGGTCGGAAGAACAACTAAATGTTTACTGGACAGATGGTTGGTACACACCTGTAATGT 488								





```
Db      876 AAGGTGTGATWAGWRAAAATTTTWKTRTTAAATGGTGATARWAAAAAAGHKGAGA 817
Qy      380 ACTGCCATGTAATTTGGAGGAAAGGACTTTAGTCAAGCTCAAGGGAATGCTGGAGAA 439
Db      816 RKGG---AGAAAAATGKGGTGWRAAAAAAAGGAGTGGAAAAAAGKRTTRTRTGT 761
Qy      440 GGCACATAATGTTTACTGGAGCAGATGGTGTGTAAACAGCCTGTATGTGCACTAACAC 499
Db      760 GTAAAAAATATGTRGTARWATAAAATTTTKKGTGATGRTAARAWAAAAATATAAAT 701
Qy      500 CAGCTGAAGAAATTAACCTAAGAGAAATAGCAGACATGAGTGGGTTACTCTACTTA 559
Db      700 AWAYGAATGAAAAAARAAAAAATGTRAWATAWATATTKATTGTGAATWGTGTTK 641
Qy      560 CTTTATAGCAATAGCAACCAAAAAAGACTATACCAAGTGTGTTCTTTTGGAAAAATGA 619
Db      640 GTKKAAAAAATAAAAAAAGKRWTAWKTAAAAWDGTTKTKTTTTRAAAAAAGGR 581
Qy      620 TTGCTTACTATTTTAACTAAAAAGAAATAAGCAGCTAGTCACCAAGAGCGGAGGCT 679
Db      580 GTGKGAAGTTRTDWAAAAAATAAAAAAAGGAGTGAADARAAGKGDAAARAD 521
Qy      680 ATTTTCTAGCAGTCTG--GCTGGAACCTAACTTTTAAAGAGCGCGAGCGCCA 737
Db      520 AGRWAAAAARATATAGTGTGKTGTGTGTAGTAAAGKTAAATWAAAAAGTTGTAKWAT 461
Qy      738 TCTAGTGCAGCACTATACACTGATGACATGCGGCCAGAAACGGTTGAAACACAGTAAC 797
Db      460 ATTAAAAAATAAAAAAAGKRWTAWKTAAAAWDGTTKTKTTTTRAAAAAAGTGT 401
Qy      798 CACTGCGCAGAACTAAGCGCGCAGAAATCAAACTAAAAAGAGTGTCTTATTAAC 857
Db      400 WTRATKTKTGTGDKAARAADAATAATTAATAAAAAAATAAAAAAATAAAAAA 341
Qy      858 TACACTTAAGAGCTGGTGCATAAAGAGTAACCTCACCAGAGGCTGGATGATGCA 917
Db      340 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 281
Qy      918 GCACAGACAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAAAACTGCTGAAAAATAC 977
Db      280 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 221
Qy      978 GCTAGAGATTTGACTACTACTAGCCAGAAACAAACAGCAATTTGACTTAATTTAGA 1037
Db      220 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 161
Qy      1038 AAAAGCTGAACACGCAACTAACCACTTTTCACTGCTGACACAGAACCTGCAGAA 1097
Db      160 NNGNANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNA 101
Qy      1098 TTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCTATTGCTGCTTTTAA 1157
Db      100 TATTASWTTTNNNNNNNANDANTATWNTANNNNNNNNNNNNNNNNNNNNNNNN 41
Qy      1158 CAGACAAGGAGGCAAAAGAAATG 1180
Db      40 NNNNANNNNNNCCCAAMTAAG 18
```

```
RESULT 14
CG753732/c
LOCUS
DEFINITION
  P048-4-G03.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
  genomic survey sequence.
ACCESSION
  CG753732
VERSION
  CG753732.1
KEYWORDS
  GSS.
SOURCE
  Pristionchus pacificus
  Pristionchus pacificus
  Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
  Neodiplogasteridae; Pristionchus.
REFERENCE
  1 (bases 1 to 1811)
```

```
AUTHORS
  Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
  Buntjer,J., van der Neulen,M. and Sommer,R.J.
TITLE
  An integrated physical and genetic map of the nematode Pristionchus
  pacificus
JOURNAL
  Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE
  22835951
PUBMED
  12884007
COMMENT
  Contact: Sommer RJ
  Evolutionary Biology
  Max-Planck-Institute for Developmental Biology
  Spemannstr. 37-39, Tuebingen D-72076, Germany
  Tel: 00497071601371
  Fax: 00497071601498
  Email: ralf.sommer@tuebingen.mpg.de
  Class: BAC ends.
FEATURES
  source
    1..1811
    /organism="Pristionchus pacificus"
    /mol_type="genomic DNA"
    /strain="California"
    /db_xref="taxon:54126"
    /clone_lib="Ppa EcoRI BAC Library"
    /notes="The library was generated by a partial digest of
    the genomic DNA with EcoRI and cloning into the BAC
    vector."
ORIGIN
  Query Match      2.3%; Score 47.2; DB 9; Length 1811;
  Best Local Similarity 34.1%; Pred. No. 1.1;
  Matches 352; Conservative 0; Mismatches 677; Indels 2; Gaps 1;
  Qy      36 AGCAACCACTGGTTAAAGAAAAAGTAACACGAGGAGTGTCTCTCTGTTTAAAAA 95
  Db      1472 ANAANNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNA 1413
  Qy      96 TGAATATGTTCACTCAATGGAAGATATCGGATGGAATAGTTACAAAAAGAGCTGCA 155
  Db      1412 AAAAAAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNA 1353
  Qy      156 GGAGGACGAGCTGAAATCTTTTCAACGAGGAGCGAAACTTCTTGGCACCACGAGGA 215
  Db      1352 NNANNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNA 1293
  Qy      216 CATGGAATGGGAAACCAACAGTGTGAATGACCAAAAGCAAGTATTCATTTTGTTC 275
  Db      1292 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1233
  Qy      276 TTTGGTTAAAAATGTTTATTTTGAAGTGTCTTACACAGAAATATATTTCTGTGATGT 335
  Db      1232 NAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1173
  Qy      336 TAATTGGTTTGTCAACATGAATGGGAAAAAGACC--AAGGCTGGCACTGCCATGTACTA 393
  Db      1172 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1113
  Qy      394 ATTGGAGGAAAGGACTTTTGTAGCTCAAGGGAATGTTGGAGAGGCAACTAAATGTT 453
  Db      1112 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1053
  Qy      454 TACTGGAGCAGATGGTTGTTAAACAGCTGTAAATGTGCACTAACACCAGCTGAAGATT 513
  Db      1052 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 993
  Qy      514 AAATAAGAGAAATAGCAGAGCAATGAGTGGGTTACTCTACTTACTTATAAGCATAAAG 573
  Db      992 ANAANNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNA 933
  Qy      574 CAAACCAAAAAAGACTATATACCAAGTGTGTTCTTTTGGAAACATGATGCTTACTATTT 633
  Db      932 NAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 873
  Qy      634 TTAACCTAAAGAAATTAAGCACTAGTCCCAAGAGCGGAGCTATTTTCTTAGCAGT 693
```

[illegible]

RESULT 15					
CNS00CFN/c					
LOCUS	CNS00CFN	933 bp	DNA	linear	GSS 04-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TtT3 end of BAC #				
	BACR25J01 of RPCI-98 library from Drosophila melanogaster (fruit				
	fly), genomic survey sequence.				

[illegible]

Search completed: January 23, 2005, 04:48:44  
Job time : 6356.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:18:38 ; Search time 99 Seconds  
(without alignments)  
2435.011 Million cell updates/sec

Title: US-10-069-056-14  
Perfect score: 3637  
Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_23Sep04:\*
- 1: Geneseq1980s:\*
  - 2: Geneseq1990s:\*
  - 3: Geneseq2000s:\*
  - 4: Geneseq2001s:\*
  - 5: Geneseq2002s:\*
  - 6: Geneseq2003as:\*
  - 7: Geneseq2003bs:\*
  - 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3627	99.7	672	4	AAy72708	Parvoviru
2	3623	99.6	672	4	AAy72702	Parvoviru
3	3620	99.5	672	4	AAy72704	Parvoviru
4	3618	99.5	672	4	AAy72710	Parvoviru
5	3618	99.5	672	4	AAy72706	Parvoviru
6	2376	65.3	690	1	AAp40306	Sequence
7	459.5	12.6	626	2	AAr85384	Barbary d
8	459.5	12.6	626	4	AAy97724	NS1 prote
9	459.5	12.6	626	5	AAE22891	Muscovy d
10	459.5	12.6	626	5	AAE28648	Muscovy d
11	459.5	12.6	626	5	AAE26944	Muscovy d
12	459.5	12.6	626	6	ABU64869	NS1 prote
13	459.5	12.6	626	6	ABU64764	Parvoviru
14	459.5	12.6	626	6	ABR43402	Muscovy d
15	459.5	12.6	626	7	ADI40288	Muscovy d
16	459	12.6	627	4	AAy97722	Rep prote
17	459	12.6	627	5	AAE22889	Barbarie
18	459	12.6	627	5	AAE28646	Barbarie
19	459	12.6	627	5	AAE26942	Barbarie
20	459	12.6	627	6	ABU64867	Rep prote
21	459	12.6	627	6	ABU64762	Parvoviru
22	459	12.6	627	6	ABR43400	Barbarie
23	459	12.6	627	7	ADI40284	Barbarie
24	453.5	12.5	550	3	AAy58164	Adeno ass
25	453.5	12.5	550	5	AAU11409	Adeno-ass

26	453.5	12.5	610	3	AAy58159	Adeno ass
27	453.5	12.5	610	4	AAy97720	Rep prote
28	453.5	12.5	610	5	AAU11404	Adeno-ass
29	453.5	12.5	610	5	AAE22887	Adeno-ass
30	453.5	12.5	610	5	AAE28644	Adeno-ass
31	453.5	12.5	610	5	AAE26940	Adeno ass
32	453.5	12.5	610	6	ABU64865	Rep prote
33	453.5	12.5	610	6	ABU64760	Adeno ass
34	453.5	12.5	610	6	ABR43398	Adeno-ass
35	453.5	12.5	610	7	ADI40280	Adeno-ass
36	453.5	12.5	610	8	ADH58895	Predeterm
37	440.5	12.1	625	7	ABR62760	Adeno ass
38	440.5	12.1	627	4	AAy97723	Rep prote
39	440.5	12.1	627	5	AAE28647	Goose par
40	440.5	12.1	627	5	AAE26943	Goose par
41	440.5	12.1	627	6	ABU64868	Rep prote
42	440.5	12.1	627	6	ABU64763	Parvoviru
43	440.5	12.1	627	6	ABR43401	Goose par
44	440.5	12.1	627	7	ADI40286	Goose par
45	434	11.9	627	5	AAE22890	Goose par

ALIGNMENTS

RESULT 1  
AAy72708  
ID AAY72708 standard; protein; 672 AA.  
XX  
AC AAY72708;  
XX  
DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX  
DE Parvovirus non-structure protein 1 (NS1) variant (T394A).  
XX  
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; mutein; variant.  
XX  
OS Parvovirus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 394 /note= "Wild type Thr substituted with Ala"  
XX  
PN EP1077260-Al.  
XX  
PD 21-FEB-2001.  
XX  
PF 13-AUG-1999; 99EP-00115161.  
XX  
PR 13-AUG-1999; 99EP-00115161.  
XX  
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
PI Nueesch J, Rommelaere J;  
XX  
PI WPI: 2001-212717/22.  
XX  
DR N-PSDB; AAD02803.  
XX  
PT Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.  
XX  
PS Claim 6; Page 25-27; 41pp: English.  
XX  
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant  
CC (T394A). The invention relates to the variants of the parvovirus non-  
CC structure protein (NS1) having a shifted equilibrium between the DNA  
CC replication and transcription activities, and the cytotoxicity activity.  
CC These variants are useful as vectors for treating tumoral diseases. The  
CC variant DNAs are useful for gene therapy. (Updated on 06-AUG-

CC	2003 to correct OS field.)	
XX		
SQ	Sequence 672 AA;	
	Query Match 99.7%; Score 3627; DB 4; Length 672;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MAGNAYSDEVLGATNWKESNQEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60	
DB	1 MAGNAYSDEVLGATNWKESNQEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60	
QY	61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNVFVQHEW 120	
DB	61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNVFVQHEW 120	
QY	121 GKDOGHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLREIAEDN 180	
DB	121 GKDOGHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLREIAEDN 180	
QY	181 EWTLLTYKHQKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240	
DB	181 EWTLLTYKHQKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240	
QY	241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRIOQTKKESIKTTLKELVHKRVTSPE 300	
DB	241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRIOQTKKESIKTTLKELVHKRVTSPE 300	
QY	301 DNMMPQDSYIENMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360	
DB	301 DNMMPQDSYIENMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360	
QY	361 TRTCRIFAFHGWNVYKCHAI CCVLRQGGKRNVLPHGPASTGKIIAQAIQAQVGNVG 420	
DB	361 TRTCRIFAFHGWNVYKCHAI CCVLRQGGKRNVLPHGPASTGKIIAQAIQAQVGNVG 420	
QY	421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOQFKAICSGQTIRIDQKKGSKQIEPTP 480	
DB	421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOQFKAICSGQTIRIDQKKGSKQIEPTP 480	
QY	481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540	
DB	481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540	
QY	541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPPTPKSTPLSQNYALTPL 600	
DB	541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPPTPKSTPLSQNYALTPL 600	
QY	601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDGLSPTWSEIEDLACFGAEP 660	
DB	601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDGLSPTWSEIEDLACFGAEP 660	
QY	661 LKXDFSEPLNLD 672	
DB	661 LKXDFSEPLNLD 672	

RESULT 2

AA72702	standard; protein; 672 AA.
ID	AA72702
XX	
AC	AA72702;
XX	
DT	06-AUG-2003 (revised)
DT	31-MAY-2001 (first entry)
XX	
DE	Parvovirus wild-type non-structure protein 1 (NS1).
XX	
KW	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW	tumoural disease; gene therapy.
XX	
OS	Parvovirus.
XX	

PN	EP1077260-A1.	
XX		
PD	21-FEB-2001.	
XX		
PF	13-AUG-1999; 99EP-00115161.	
XX		
PR	13-AUG-1999; 99EP-00115161.	
XX		
PA	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	
XX		
PI	Nueesch J, Rommelaere J;	
XX		
DR	WPI; 2001-212717/22.	
DR	N-PSDB; AAD02797.	
XX		
PT	Novel parvovirus non-structure protein variant, useful for treating	
PT	tumoral diseases, has a shifted equilibrium between DNA replication and	
PT	transcription activities, and cytotoxic activity.	
XX		
PS	Disclosure; Fig 1; 41pp; English.	
XX		
CC	The present sequence is a parvovirus wild-type non-structure protein 1	
CC	(NS1). The present invention relates to the variants of the parvovirus	
CC	non-structure protein (NS1) having a shifted equilibrium between the DNA	
CC	replication and transcription activities, and the cytotoxicity activity.	
CC	These variants are useful as toxins for treating tumoural diseases. The	
CC	variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-	
CC	2003 to correct OS field.)	
XX		
SQ	Sequence 672 AA;	
	Query Match 99.6%; Score 3623; DB 4; Length 672;	
	Best Local Similarity 99.7%; Pred. No. 0;	
	Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 MAGNAYSDEVLGATNWKESNQEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60	
DB	1 MAGNAYSDEVLGATNWKESNQEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60	
QY	61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNVFVQHEW 120	
DB	61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNVFVQHEW 120	
QY	121 GKDOGHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLREIAEDN 180	
DB	121 GKDOGHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLREIAEDN 180	
QY	181 EWTLLTYKHQKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240	
DB	181 EWTLLTYKHQKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240	
QY	241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRIOQTKKESIKTTLKELVHKRVTSPE 300	
DB	241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRIOQTKKESIKTTLKELVHKRVTSPE 300	
QY	301 DNMMPQDSYIENMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360	
DB	301 DNMMPQDSYIENMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360	
QY	361 TRTCRIFAFHGWNVYKCHAI CCVLRQGGKRNVLPHGPASTGKIIAQAIQAQVGNVG 420	
DB	361 TRTCRIFAFHGWNVYKCHAI CCVLRQGGKRNVLPHGPASTGKIIAQAIQAQVGNVG 420	
QY	421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOQFKAICSGQTIRIDQKKGSKQIEPTP 480	
DB	421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOQFKAICSGQTIRIDQKKGSKQIEPTP 480	
QY	481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540	
DB	481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540	
QY	541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPPTPKSTPLSQNYALTPL 600	
DB	541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPPTPKSTPLSQNYALTPL 600	
QY	601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDGLSPTWSEIEDLACFGAEP 660	
DB	601 ASDLEDLALPWSPTNTPVAGTAETQNTGEAGSKACQDGLSPTWSEIEDLACFGAEP 660	
QY	661 LKXDFSEPLNLD 672	
DB	661 LKXDFSEPLNLD 672	

Db	541	NGYOSTMASYCAKWKGVDPWSENWAEKPVPTPINLLGSARSFPTTPKSTPLSQNYALTPL	600
Qy	601	ASDLEDLALPWSPTNPTVAGTAETONTGEAGSKACQDQQLSPTWSEIEEDLRACFGAEP	660
Db	601	ASDLEDLALPWSPTNPTVAGTAETONTGEAGSKACQDQQLSPTWSEIEEDLRACFGAEP	660
Qy	661	LKKDFSEPLNLD 672	
Db	661	LKKDFSEPLNLD 672	
RESULT 3			
ID	AAAY72704	standard; protein; 672 AA.	
XX	AC	AAAY72704;	
XX	AC		
DT	06-AUG-2003	(revised)	
DT	31-MAY-2001	(first entry)	
XX	XX	Parvovirus non-structure protein 1 (NS1) variant (S283A).	
DE	XX	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;	
KW	XX	tumoural disease; gene therapy; mutant; mutein; variant.	
XX	OS	Parvovirus.	
OS	OS	Synthetic.	
XX	XX		
FT	Key	Location/Qualifiers	
FT	Misc-difference	283	
XX	XX	/note= "Wild type Ser substituted with Ala"	
XX	XX	Ep1077260-A1.	
XX	XX	21-FEB-2001.	
XX	XX	13-AUG-1999; 99EP-00115161.	
XX	XX	13-AUG-1999; 99EP-00115161.	
XX	XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	
XX	XX	Nueesch J, Rommelaere J;	
PI	PI	WPI: 2001-212717/22.	
DR	DR	N-PSDB; AAD02799.	
XX	XX	Novel parvovirus non-structure protein variant, useful for treating	
PT	PT	tumoral diseases, has a shifted equilibrium between DNA replication and	
PT	PT	transcription activities, and cytotoxic activity.	
XX	XX	Claim 6; Page 14-16; 41pp; English.	
PS	PS	The present sequence is parvovirus non-structure protein 1 (NS1) variant	
CC	CC	(S283A). The invention relates to the variants of the parvovirus non-	
CC	CC	structure protein (NS1) having a shifted equilibrium between the DNA	
CC	CC	replication and transcription activities, and the cytotoxicity activity.	
CC	CC	These variants are useful as toxins for treating tumoural diseases. The	
CC	CC	variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-	
CC	CC	2003 to correct OS field.)	
XX	XX	Sequence 672 AA;	
XX	XX	Query Match 99.5%; Score 3620; DB 4; Length 672;	
XX	XX	Best Local Similarity 99.6%; Pred. No. 0;	
XX	XX	Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
Qy	1	MAGNAYSDEVLGATNWLKESKNOEVSFVKENYVQLNGKIDGWSYKKEQLQDELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKESKNOEVSFVKENYVQLNGKIDGWSYKKEQLQDELKSLQ	60
Qy	61	RGAEATTWQSDMEWETTVDKTKQVIFDLSLVKKCLFEVLNTKNIPFGDVNFWQHEW	120

```
XX Nueesch J, Rommelaere J;
PI WPI; 2001-212717/22.
DR N-PSDB; AAD02805.
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX Claim 6; Page 30-32; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX SQ Sequence 672 AA;

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWDSQSEDMWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNIFPGDVNWFVQHEW 120
DB 61 RGAETTWDSQSEDMWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNIFPGDVNWFVQHEW 120
QY 121 GKQGWCHVLIGKDSQAGQKWRRLQNVYWSRWLVLTACNVQLTPAERIKLRETAEDN 180
DB 121 GKQGWCHVLIGKDSQAGQKWRRLQNVYWSRWLVLTACNVQLTPAERIKLRETAEDN 180
QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
DB 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240

QY 241 KEGERHLVSKLYTDDMRPEVETVTTAQTETKGRIGTQKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPEVETVTTAQTETKGRIGTQKKEVSIKTTLKELVHKRVTSPE 300
QY 301 DWMWQSDSYEMMAQPGGNNLLKNTLEICTLTLARTKTAFDLLEKAETSKLTNSLDP 360
DB 301 DWMWQSDSYEMMAQPGGNNLLKNTLEICTLTLARTKTAFDLLEKAETSKLTNSLDP 360
QY 361 TRTCRIFAFHGWNVYKCHVCAICCVLNRQGGKRNVLPHFGPASTGKSIIAQIAQAVGNVG 420
DB 361 TRTCRIFAFHGWNVYKCHVCAICCVLNRQGGKRNVLPHFGPASTGKSIIAQIAQAVGNVG 420
QY 421 CYNAANVFPNDCTNKNLWBEAGNFGQVNOFKAICSGQITRIDQKKGSKQIEPTP 480
DB 421 CYNAANVFPNDCTNKNLWBEAGNFGQVNOFKAICSGQITRIDQKKGSKQIEPTP 480
QY 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDLVDKNEWPMICAWLVK 540
DB 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDLVDKNEWPMICAWLVK 540
QY 541 NGYQSTMASYCAKMGKVPDSENWAEKVPPTPINLLGSARSPPTPKSTPLSQNYALTPL 600
DB 541 NGYQSTMASYCAKMGKVPDSENWAEKVPPTPINLLGSARSPPTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPWPSTPNTPVAGTAETQNTGAGSKACODGQLSTWSEIBEDLRACGAEP 660
DB 601 ASDLEDLALPWPSTPNTPVAGTAETQNTGAGSKACODGQLSTWSEIBEDLRACGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672
```

```
RESULT 5
AAY72706
ID AAY72706 standard; protein; 672 AA.
XX
XX AAY72706;
AC AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoral disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 363
FT /note= "Wild type Thr substituted with Ala"
FT
XX EP1077260-Al.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
PR (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
DR N-PSDB; AAD02801.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX Claim 6; Page 19-21; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T363A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX SQ Sequence 672 AA;

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWDSQSEDMWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNIFPGDVNWFVQHEW 120
DB 61 RGAETTWDSQSEDMWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNIFPGDVNWFVQHEW 120
QY 121 GKQGWCHVLIGKDSQAGQKWRRLQNVYWSRWLVLTACNVQLTPAERIKLRETAEDN 180
DB 121 GKQGWCHVLIGKDSQAGQKWRRLQNVYWSRWLVLTACNVQLTPAERIKLRETAEDN 180
QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
DB 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
```

QY 241 KGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKKEVSIKTTLKELVHKRVTSPE 300  
 DB 241 KGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKKEVSIKTTLKELVHKRVTSPE 300  
 QY 301 DWMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDILILEKAETSKLTNFSLPD 360  
 DB 301 DWMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDILILEKAETSKLTNFSLPD 360  
 QY 361 TRTCRIFAFHGMNVYKCHACCVLNRQGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG 420  
 DB 361 TRTCRIFAFHGMNVYKCHACCVLNRQGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG 420  
 QY 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKGGSKQIEPTP 480  
 DB 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKGGSKQIEPTP 480  
 QY 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540  
 DB 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540  
 QY 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600  
 DB 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600  
 QY 601 ASDLEDLALPWPSTPNTPVAGTAETQNTGEGAGSKACODGQLSPTWSEIIBDLRACFGAEP 660  
 DB 601 ASDLEDLALPWPSTPNTPVAGTAETQNTGEGAGSKACODGQLSPTWSEIIBDLRACFGAEP 660  
 QY 661 LKXDFSEPLNLD 672  
 DB 661 LKXDFSEPLNLD 672

## RESULT 6

AA040306  
 ID AAP40306 standard; protein; 690 AA.

AC  
 AC AAP40306;

DT 24-OCT-2003 (revised)  
 DT 12-JAN-1992 (first entry)

XX Sequence encoded by the double-stranded replicative form DNA of porcine parvovirus.

KW Protein envelope; immunogen; vaccine; antigen; epitope.

OS Porcine parvovirus; NADL-2 virulent strain.

PN WO8402847-A.

XX 02-AUG-1984.

XX 19-JAN-1984; 84WO-US0000063.

XX 19-JAN-1983; 83US-00459203.

PR 06-JAN-1984; 84US-00567968.

XX (AMGE-) AMGEN.

XX Fox GW;

XX WPI; 1984-201354/32.

DR N-PSDB; AAN40252.

XX Polypeptide obtd. by recombinant DNA methods - for vaccination against parvovirus infections in man and animals.

XX Claim 6; Table II, Page 33-49; 80pp; English.

XX The inventors claim an immunologically active polypeptide for the development of vaccinal immunity against parvovirus infection. Also

CC

CC claimed are DNA sequences wholly or partly duplicative of defined sequences. The polypeptides are used in vaccines for conferring protection against parvovirus infections in man and animals. (Updated on CC 24-OCT-2003 to standardise OS field)

XX SQ Sequence 690 AA;

Query Match 65.3%; Score 2376; DB 1; Length 690;

Best Local Similarity 67.8%; Pred. No. 2.3e-221;

Matches 449; Conservative 70; Mismatches 127; Indels 16; Gaps 6;

QY 2 AGNAYDSEVLGATNWLKESNQEVFVFNENVNQNGKIDGWSYKKEIQEDELKSLQR 61  
 DB 5 AGNTYSEVLKATNWLQDNAQKEAFSVFTQKVNLRNGEIAWNNYKNTDTDAEMINLQR 64

QY 62 GAETTTWQSDSEMEWETTVTDEMTKKQVFI FDSL VKCLFEVLNNTKNI FPGDVNMFVQHEWG 121  
 DB 65 GAETSDQATDMEWESEIDSLTKRQVLI FDSL VKCLFEGLIKQLKSPSCYWFQLQHEHG 124

QY 122 KDQGWCHVLIGKDFSOAQGWRRQLNVYNSRWLVTAQNVLTFAERIKLREIAEDNE 181  
 DB 125 QDTGYHCHVLLGGKGLQAMGKWLRLQNLNLSRWLNQCKVPLTPVERIKLRELAEDGE 184

QY 182 WTLTLLTKKQTKDYTKCVLFCNMIAYYFLTKKKISTSPRDGGYFLSSDSQWKTWFLK 241  
 DB 185 WSLTYTHKQTKQYTKMTHFGNMIAYYFLNKKRKITE--REHYVLLSSDSGFWNLFK 242

QY 242 EGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKKEVSIKTTLKELVHKRVTSPE 301  
 DB 243 EGERHLVSHLFTANKPETVETVTTAQTKEGRIOQTKKEVSIKTTIRDLVNRKCTSI EG 302

QY 302 WMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDILILEKAETSKLTNFSLPDT 361  
 DB 303 WSMTPDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDILILEKAKPSMLPTFNISNT 362

QY 362 RTCRIFAFHGMNVYKCHACCVL--NRQGGKRNVLPHGPASTGKSIIAQAIAQAVGNV 419  
 DB 363 RTCKIFSMHNNYIKVCHAITCVLKQTRKKKNSISCHGPASTGKSIIAQHIANLVGNV 422

QY 420 GCYNAANVFPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKGGSKQIEPT 479  
 DB 423 GCYNAANVFPNDCTNKNLIWVEAGNFSNQVQFKAI CSQTIRIDQKGGSKQIEPT 482

QY 480 PVIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLV 539  
 DB 483 PVIMTTNEDITKVRIGCEERPEHTQPIRDRMLNHLTRKLPGLGLEETWEPFLICAWLV 542

QY 540 KNGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPT 599  
 DB 543 KNGYQATMASYMHWNVPDSEKLBEPKMHSPINTPTDSQIS-TSVKTSPADINYAATP 601

QY 600 LASDLE-DLALPWPSTPNTPVAGTAETQNTGEGAGSKACODGQL---SPTWSEIEEDLRAC 655  
 DB 602 IQEDLDLALPWPSEPTPTFTNLHLTPP-----PDSAIRTPSPWSEIETDIRAC 654

QY 656 FG 657  
 DB 655 FG 656

## RESULT 7

AAR85384

ID AAR85384 standard; protein; 626 AA.

XX AAR85384;

XX 27-AUG-2003 (revised)

DT 09-MAY-1996 (first entry)

XX Barbary duck parvovirus fragment ORF1 encoded non-structural protein.

XX Barbary duck; parvovirus; open reading frame; non-structural; VP1; VP2;

KW liver extract; virus particle; vector; host cell; attenuated; geese.

```
XX OS Barbarie duck parvovirus.
XX Key Location/Qualifiers
XX FT Misc-difference 213 /note= "encoded by GTA"
XX PT
XX PR FR2718151-A1.
XX PN
XX PD 06-OCT-1995.
XX PF 03-MAR-1995; 95FR-00002515.
XX PR 03-MAR-1994; 94HU-00000634.
XX PA (SNFI ) PHYLAXIA-SANOFI OLTOANYACTERMELO RT.
XX PI Kisari J, Zadori Z, Erdei J, Nagy J;
XX DR WPI; 1995-346564/45.
XX DR N-PSDB; AAT01190.
XX CC Nucleic acid sequence from Barbary duck parvovirus - related vectors,
XX PT transformed cells, proteins etc., useful diagnostically and in protective
XX PT vaccines.
XX PS Claim 2; Fig 3; 39pp; French.
XX CC The sequence of the non-structural protein encoded by the open reading
XX CC frame (ORF) 1 contained in a fragment from the Barbary duck parvovirus
XX CC strain FM. The fragment contains 3 ORFs, with ORF3 being internal to
XX CC ORF2, encoding a non-structural, the VP1 and VP2 proteins respectively.
XX CC The virus was isolated from a 13 day old duck that had died of Barbary
XX CC duck parvovirus infection, by incubating liver extracts injected into
XX CC embryonated duck eggs. The DNA was isolated from the resulting 20-25 nm
XX CC viral particles. The DNA sequences encoding the non-structural, VP1
XX CC (AAR85385) or VP2 (AAR85386) proteins can be used to construct viral
XX CC vectors, host cells or attenuated parvoviruses for use as vaccines to
XX CC protect Barbary ducks and geese. (Updated on 27-AUG-2003 to correct OS
XX CC field.)
XX SQ Sequence 626 AA;

Query Match 12.6%; Score 459.5; DB 2; Length 626;
Best Local Similarity 25.3%; Pred. No. 9.5e-35;
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

QY 75 WEIT-----VDENTKKQVFIEDSLVKKCLFEVLTAKNIF-----PGDVNWFVQHWGK 122
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 45 WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQRMNQFNQDETDFPFQLEEGS 93
QY 123 DOGWHCHVLGGKDFSOAQCKWRRQLNVYWSRLVTACNVQLTPAERIK---LRRIAD 179
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 94 EY-IHLHAVCPGECRSEVLGRY-----MSQIKDSLRLDVYEG 129
QY 180 N-----EWVTLTYKHQTKDYKCVLFGNMIAYYFLTKKXISTPPRDGGYFLSSDSG 234
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 130 KQVKIPDWFSITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179
QY 235 WKTNFLKEGRHLVSKLYTDDMPETVETVTTTAQETKRGRIQTKVEVSKITLKLVLHK 294
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 180 TAAALCLQKQKQELLDAPQSEMNVAVQEDQAAPL-----ISNRAAKNYSNLVDWLIEM 235
QY 295 RVTSPEDDMMWQPSYIEMMAQPGGENLLKNLTLEICTLTARTKTAFLDILEKAETSKLT 354
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 236 GIYSEKQWLTENKESYSPATSSNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295
QY 355 NFSLPDTRTCRIFAFHGWNVYKCHAICCVLNRQGGKRNVAFLPHGPASTGKSIQAQIAQ 414
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 296 K-----NRIYQILKLNNYPQYGVSVLCGWVKFEFNKRNALWLYGPATTKTIAEIAH 350
QY 415 AVGNVCYNAANYNFPENDCTKNLWVEAGNFGQVQVQFKACISQGITRIDOKSGK 474
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 AVPFYGCNVNNTNENFPFNDVCVDMXLIWEEGKMTNKVVSAILGSAVRVDQCKGKSV 410
QY 475 QIEPTVIMTNNENITVWRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNE---- 530
DB ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 411 CIEPTVITSTNDCMIVDGNSTTWHRIPLEERMFQIVLSKLEGNFKISKKEVKEF 470
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKMGKVPDWSN-----NAEPK----- 568
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPSEPPKIMAPPTREELEEI 524
QY 569 -----VPTPINLIGSARSPTTPKSTPLSQNYALTPLASDIEDLALPEWSTPNT 617
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 525 LRASPELFASVAPLP-----SSDTPSKRKKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETONTGAGSKAC-QDGQLSPTWSEIEDLRACFAGBPLKKDFSE 667
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDWNE 612

RESULT 8
AAY97724
ID AAY97724 standard; protein; 626 AA.
XX AC AAY97724;
XX DT 19-JUN-2001 (first entry)
XX DE NS1 protein sequence.
XX KW Fusion nucleic acid library; Rep protein; tumour cell; apoptosis;
XX KW nucleic acid modification enzyme; cell death; decreased cell growth;
XX KW protein-protein interaction detection; cell division; cancer therapy;
XX KW protein drug discovery; pharmacogenetics; NS1 protein; ds.
XX OS Muscovy duck parvovirus.
XX PN WO200114539-A2.
XX PD 01-MAR-2001.
XX PF 18-AUG-2000; 2000WO-US022906.
XX PR 20-AUG-1999; 99US-0150004P.
XX PR 02-JUN-2000; 2000US-0209130P.
XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX PI Li M;
XX WPI; 2001-218443/22.
XX DR N-PSDB; AAA91314.
XX PT New library of fusion nucleic acids each encoding a Rep protein
XX PT recognized by a nucleic acid modification enzyme and a candidate protein,
XX PT useful for detecting protein-protein interactions, protein drug discovery
XX PT or pharmacogenetics.
XX PS Disclosure; Fig 29; 106pp; English.
XX CC This sequence is the muscovy duck parvovirus NS1 protein. The invention
XX CC relates to a library of fusion nucleic acids, each encoding a Rep
XX CC protein, a candidate protein, a presentation structure, a targeting
XX CC sequence or a label. The Rep protein is a nucleic acid modification
XX CC enzyme. The random or directed libraries (including the cDNA libraries)
XX CC can be introduced into any tumour cell, and peptides identified which by
XX CC themselves induce apoptosis, cell death, loss of cell division or
XX CC decreased cell growth. The methods and compositions may also be used to
XX CC detect protein-protein interactions, protein drug discovery, particularly
XX CC for protein drugs that interact with targets on cell surfaces, to
XX CC discover DNA or nucleic acid binding proteins, using nucleic acids as
XX CC targets, to screen for nucleic acid modification enzymes with decreased
XX CC toxicity for the host cells, to identify or generate Rep proteins with
XX CC decreased toxicity, improved enzyme attachment sequences for use in
```







XX PD 29-AUG-2002.  
XX PF 14-DEC-2001; 2001WO-US049058.  
XX PR 14-DEC-2000; 2000US-0256163P.  
XX PA (XENC-) XENCOR INC.  
XX PI Li M, Liu Y;  
XX DR WPI; 2002-667068/71.  
XX DR N-PSDB; AAD4604.  
XX PT New library of prokaryotic pBT-24a expression vectors, host cells or  
XX PT nucleic acid/protein conjugates, useful for screening candidate proteins  
XX PT and their nucleic acids or modification enzymes for pharmacogenetic  
XX PT analysis.  
XX PS Disclosure; Fig 29; 127pp; English.  
XX CC The invention relates to methods and compositions for the construction of  
XX CC prokaryotic libraries expressing candidate proteins and the use of these  
XX CC libraries to identify candidate proteins and the nucleic acids encoding  
XX CC them. The invention provides a library of prokaryotic pBT-24a vectors  
XX CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a  
XX CC nucleic acid modification (NAM) enzyme or a candidate protein, or a  
XX CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or  
XX CC by the NAM enzyme. The library is used for identifying candidate proteins  
XX CC and nucleic acids encoding these proteins, in screening for NAM enzymes  
XX CC with decreased toxicity for the host cells, or in identifying novel or  
XX CC improved EASs, which may be used for understanding cellular processes or  
XX CC any subsequent therapeutic or toxic activities. The nucleic acid/protein  
XX CC (NAP) conjugates are useful in diagnostic assays and in research  
XX CC including clinical pharmacology, functional genomics, pharmacogenomics,  
XX CC agricultural chemicals, environmental safety assessment, chemical sensor,  
XX CC nutrient biology, cosmetic research or enzymology. These may also be used  
XX CC in vitro screening techniques and in assays with target molecules. The  
XX CC present sequence is Muscovy duck parvovirus NS1 protein used in the  
XX CC invention  
XX SQ Sequence 626 AA;  
Query Match 12.6%; Score 459.5; DB 5; Length 626;  
Best Local Similarity 25.3%; Pred. No. 9.5e-35;  
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;  
75 WETT-----VDEMTKKQVIFDLSVKKCLFEVLTKNIP-----PGDVNWFVQHEWKG 122  
45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRWNQFNQDETDFPFQLEEGS 93  
123 DQGWCHVLIGKDFSAQGWKWRQLNVYWSRWLVYACNVLTAPRIK---LRETAED 179  
94 EY-IHLHAVCGEGRSEFVLGRY-----MSQIKDSILRDVYEG 129  
180 N-----EWVTLATYKHQTKDYKCVLFGNMIAYFLTKKISTSPRDGCVFLGSDSG 234  
130 KQVKIPDWFSTIKTRGGQNKVTAA-----YILHVLIPKQ-----PELQWAFNTNPLF 179  
235 WKTNFKEGERHLVSKLYTDDMPETVETVTTTAQETKRGRIOTKKEVSIKTLKELVHK 294  
180 TAAALCLQKQELLDAFOESEMNAVQEDQASTAAPL-----ISNRAKNYSNLVDWLIEM 235  
295 RVTSPEWMMQPDYSIEMMAQPGENLLKNTLEICTLTARPKTAPDLILEKAEYSKLT 354  
236 GITSEKQWLTENKESYRSFQATSSNNQVRAALENARAEMLLTKATDYILIGKDPVLDDIT 295  
355 NFSLPDTRTCRIFAFHGMWYVVKVCHAJCCVNLNROGGKRNVALPHGPASTGKSIIAQIAQ 414  
296 K-----NRIYQILKLNYNFQYVGSVLCGWVKREFNKNALWLYGPATTKTIAEIAH 350  
415 AVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQOVNOFFKAICSGQTIRIDQKGGSK 474

Db 351 AVFPYGCWNTNENFFNDVCVKNLIWEEGKTNKVESAKAILGGSVAVDOKCKGVS 410  
Qy 475 QIEPTVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNE---- 530  
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568  
Db 471 FKWANDNLVFWSEFKVPTNEQTKLTE-----VPERANEPSEPPKIMAPPTREELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALBEPWSPNT 617  
Db 525 LRASPELFRASVAPLP-----SSPTSPKPKTKRGEYQVRCAMHSL-DNSMNVFCELEC 576  
Qy 618 PVAGTAETQNTGAGSKAC-QDQQLSPTWSEIEEDLRACFGAPLKKDFSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 12  
ABU64869  
ID ABU64869 standard; protein; 626 AA.  
XX AC ABU64869;  
XX DT 14-MAY-2003 (first entry)  
XX DE NS1 protein sequence from Muscovy duck parvovirus.  
XX KW Rep protein; capture probe; expression vector;  
XX KW nucleic acid protein conjugate; NAP; enzyme attachment sequence; EAS;  
XX KW biochip; gene expression profiling; mutation detection; Rep68; Rep78;  
XX KW nonstructural protein; NS1; major coat protein; U94.  
XX OS Muscovy duck parvovirus.  
XX PN US2002172968-A1.  
XX PD 21-NOV-2002.  
XX PF 19-FEB-2002; 2002US-00080376.  
XX PR 22-FEB-2001; 2001US-00792630.  
XX PA (LIUH/) LIU H.  
XX PA (DAHI/) DAHIYAT B I.  
XX PA (LIHM/) LI M.  
XX LIU H, Dahiyat BI, Li M;  
XX WPI; 2003-310986/30.  
XX DR N-PSDB; ABX96673.  
XX PT New composition comprising a substrate consisting of an array of capture  
XX PT probes hybridized to an expression vector or to a nucleic acid protein  
XX PT conjugate, useful for diagnostic test, gene expression profiling or  
XX PT mutation detection.  
XX PS Disclosure; Fig 29; 125pp; English.  
XX CC The invention relates to a composition comprising a substrate comprising  
XX CC an array of capture probes hybridized to an expression vector or to a  
XX CC nucleic acid protein conjugate. The capture probes are hybridized to an  
XX CC expression vector or to a nucleic acid protein (NAP) conjugate. The  
XX CC vector comprises: (a) a fusion nucleic acid; (b) a capture sequence; and  
XX CC (c) an enzyme attachment sequence (EAS). The NAP conjugate comprises: (a)  
XX CC a fusion polypeptide; and (b) an expression vector. The fusion nucleic  
XX CC acid comprises a nucleic acid encoding the NAP enzyme or candidate  
XX CC protein. The fusion polypeptide comprises a Rep and candidate protein.  
XX CC The EAS and NAP enzyme are covalently attached. Also included are  
XX CC detecting the presence of a target analyte in a sample, making biochips,  
XX CC and making NAP conjugates. The composition is useful for diagnostic





XX 22-FEB-2002; 2002WO-US008023.  
 PF 22-FEB-2001; 2001US-0270781P.  
 PR (XENC-) XENCOR INC.  
 XX Dahiyat B, Li M;  
 XX WPI; 2003-040516/03.  
 DR N-PSDB; ADI40289.  
 XX  
 PT Creating protein-binding profile of test compound, to create  
 PT toxicological profile indicative of compound's toxicity in vivo, involves  
 PT screening expression library to identify proteins that interact with  
 PT compound.  
 XX  
 PS Disclosure; SEQ ID NO 35; 109pp; English.  
 XX  
 CC The invention relates to a novel method for creating a protein-binding  
 CC profile of test compound (T) comprising contacting (T) with library of  
 CC nucleic acid (NA)/protein conjugates (I), where each (I) has a fusion  
 CC protein with NA modifying enzyme (II), and candidate protein (III), and  
 CC an expression vector with fusion NA comprising coding sequences for (II)  
 CC and (III), and enzyme attachment sequence (EAS). The method of the  
 CC invention is useful for carrying out species-specific toxicology tests,  
 CC differential organ interaction tests, developmental stage-specific  
 CC toxicity tests, and individualised toxicity tests. The method allows for  
 CC the creation of a protein-binding profile of the test compound e.g., any  
 CC synthetic or natural compound including an organic or inorganic compound,  
 CC a peptide or nucleic acid, a metabolite, a drug derivative, or a chemical  
 CC entity. The compound can be drug, drug candidate, or an ingredient in  
 CC human consumable (e.g., food, textile, cosmetics, flavours, fragrances,  
 CC emulsifiers, surfactants, and detergents); compounds that in come in  
 CC contact with humans and other animals (e.g., pets and farm stock) such  
 CC as pesticides, fertilizers, feed additives, antibiotics, herbicides,  
 CC fungicides, polymer additives, and environmental proteins. These profiles  
 CC can then be used to evaluate and predict toxicity and other biological  
 CC activities of the test compounds. The toxicity profiling information can  
 CC be used to predict toxicity of a compound in a different species, e.g.,  
 CC to extrapolate the toxicity effects of a compound from one species to  
 CC another. The information can also be used to predict the toxicity of a  
 CC compound in individuals of the same species that are under different  
 CC physiological conditions (e.g., age, sex, and/or disease states, e.g., to  
 CC identify individuals particularly susceptible to the toxicity.  
 CC Preferably, the method is used to assess the safety of a candidate drug  
 CC prior to clinical trial, and to improve clinical trials by allowing  
 CC determination of toxic or unanticipated responses in humans early in a  
 CC clinical trial to avert tissue toxicity. The methods also helps evaluate  
 CC and predict toxicity of chemicals in environmental or occupational  
 CC settings such as in manufacturing and agriculture. The present sequence  
 CC is used in the exemplification of the invention. Note: The sequences  
 CC given in the figures of the specification (ADI40260-ADI40307) are wrongly  
 CC labelled in the figure legends. Protein sequences are stated as being  
 CC nucleic acids, and vice versa.  
 XX  
 SQ Sequence 626 AA;

Query Match 12.6%; Score 459.5; DB 7; Length 626;  
 Best Local Similarity 25.3%; Pred. No. 9.5e-35;  
 Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

QY 75 WETT-----VDMETKKQVFIDSLVKKGLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
 DB 45 WEPTGIWNMEHVNLPMTLADKI-----KNIFIQRWNQFNQDETDFPQLEGS 93  
 QY 123 DQGHCHVLIGGKDFSOAGKWRRLNVLVWRSRLVTACNVQLTPAERIK---LREIAD 179  
 DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSLRLDYEG 129  
 QY 180 N-----EWVILLTYKHQTKKDYKCVLFGNVIAYFLTKKISTSPPRDGGVFLSSDSG 234  
 DB 130 KQVKIPDFWSITKTRGGQNTVTAA-----YILHLIPKQ-----PELOWAFNTNPLF 179

QY 235 WKTNFKEGERHLVSKLYTDDMRPPTVTTVTAQETKRGRIOTKVEYSIKTTKLVLHK 294  
 DB 180 TAAALCLQKQBELLDAFQSEMNNAVQEDQASTAPL-----ISNRAAKNYSNLVDWLIEM 235  
 QY 295 RVTSPEDWMMQPDYSIEMMAOPGGENLLKNLTLEICTLTARTKTAFDLILEKAETSKLT 354  
 DB 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295  
 QY 355 NFSLPDTRTCRIFAFHGNVYKCHAI CCVLNRQGGKENAVLFHGPASTGKSIIAQAIQA 414  
 DB 296 K-----NRIYQILKLNYPQYVGVLCGWVKBEPFNKNAIWLGYPATTKTINIAEIAH 350  
 QY 415 AVGNVGCYNAANVNPFPNDCTNKNLIWVEHAGNFQQOVNQFKAI CSQTIRIDQKGKSK 474  
 DB 351 AVPFYGCNVWTNENFPFND CVDKMLIWWEEGKNTKVVESAKAILGGSRAVRVDQCKGVS 410  
 QY 475 QIEPTFVIMTNTENITVVRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDKNE---- 530  
 DB 411 CIEPTFVIITNDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLBNFGKISKEVKEF 470  
 QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568  
 DB 471 FKWANDNLVPVVSSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELEBI 524  
 QY 569 -----VPTPINLIGSARSPTTPKSTPLSQNTYALTPLASDLEDLALPEHSTPNT 617  
 DB 525 LRASPELFASVAPLP-----SSPDTSPKRXKTRGEYQVRCAMHSL-DNSMNVFECLEEC 576  
 QY 618 PVAGTAETONTGEAGSKAC-QDQQLSPTWSETEEDLRACFGAEPLKDPSE 667  
 DB 577 ERANPPEFOSLGE---NFCNQHG-----WYD-----CAFNCNELKDDMNE 612

Search completed: January 22, 2005, 03:35:25  
 Job time : 102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:50 ; Search time 25 Seconds  
(without alignments)  
1782.628 Million cell updates/sec

Title: US-10-069-056-14  
Perfect score: 3637  
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432.5	11.9	536	4	US-09-532-594B-10
2	432.5	11.9	623	4	US-09-532-594B-2
3	432.5	11.9	623	4	US-09-532-594B-11
4	428.5	11.8	546	4	US-09-807-802A-7
5	428.5	11.8	623	4	US-09-807-802A-2
6	428.5	11.8	623	4	US-09-807-802A-5
7	425.5	11.7	312	4	US-09-532-594B-8
8	425.5	11.7	399	4	US-09-532-594B-9
9	420.5	11.6	322	4	US-09-807-802A-11
10	420.5	11.6	399	4	US-09-807-802A-9
11	139	3.8	276	3	US-09-171-461-3
12	139	3.8	276	3	US-09-970-711-3
13	124.5	3.4	647	3	US-09-300-909-17
14	114	3.1	564	3	US-09-134-001C-2897
15	113	3.1	589	4	US-09-107-532A-6086
16	104	2.9	1170	4	US-09-638-524A-2
17	101	2.8	715	4	US-09-248-796A-20602
18	101	2.8	1120	4	US-09-792-024-95
19	100.5	2.8	2409	6	5180808-2
20	100	2.7	481	4	US-09-252-991A-30005
21	100	2.7	605	4	US-09-107-532A-6077
22	100	2.7	1503	4	US-09-677-046A-6
23	100	2.7	1509	4	US-09-677-046A-2
24	99.5	2.7	657	3	US-09-300-909-16
25	99	2.7	919	4	US-09-248-796A-19408
26	97.5	2.7	556	4	US-09-248-796A-18763
27	97.5	2.7	3248	1	US-08-353-700-1

28	97.5	2.7	3248	5	PCT-US95-16216-1	Sequence 1, Appli
29	96.5	2.7	557	4	US-09-489-039A-11233	Sequence 11233, A
30	96	2.6	291	4	US-09-107-532A-6390	Sequence 6390, Ap
31	95.5	2.6	1060	4	US-09-248-796A-16624	Sequence 16624, A
32	95.5	2.6	3898	3	US-08-750-717-2	Sequence 2, Appli
33	95	2.6	590	4	US-09-134-000C-4755	Sequence 4755, Ap
34	94	2.6	370	4	US-09-583-110-3405	Sequence 3405, Ap
35	94	2.6	471	4	US-09-134-000C-6228	Sequence 6228, Ap
36	94	2.6	605	2	US-08-472-666-1	Sequence 1, Appli
37	94	2.6	605	5	PCT-US96-07615-1	Sequence 1, Appli
38	94	2.6	1115	2	US-08-568-459A-2	Sequence 2, Appli
39	94	2.6	1115	2	US-08-487-826B-2	Sequence 2, Appli
40	94	2.6	1115	3	US-09-210-288-2	Sequence 2, Appli
41	94	2.6	1115	6	5198347-6	Patent No. 5198347
42	94	2.6	1307	1	US-08-395-246C-2	Sequence 2, Appli
43	94	2.6	1343	4	US-09-270-767-31792	Sequence 31792, A
44	94	2.6	1343	4	US-09-270-767-47009	Sequence 47009, A
45	94	2.6	1477	4	US-09-206-942-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1

US-09-532-594B-10  
; Sequence 10, Application US/09532594B  
; Patent No. 6468524  
; GENERAL INFORMATION:  
; APPLICANT: Chorini, John A.  
; APPLICANT: Kotin, Robert M.  
; APPLICANT: Safer, Brian  
; APPLICANT: Davidson, Beverly  
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF  
; FILE REFERENCE: 14014.0252U2  
; CURRENT APPLICATION NUMBER: US/09/532,594B  
; CURRENT FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 536  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =  
; OTHER INFORMATION: synthetic construct  
; NAME/KEY: misc feature  
; OTHER INFORMATION: AAV4 Rep protein 68  
US-09-532-594B-10

Query Match	11.9%	Score	432.5	DB	4	Length	536
Best Local Similarity	27.5%	Pred.	No. 8.4e-36				
Mismatches	156	Conservative	75	Mismatches	233	Indels	103
Gaps	19						
Qy	63	AETTQSDSEMEWETTVDEMTKKQVFIFDSLVKCLFEVLNTKNIIPGDVNWVFQHEWGK	122				
Db	31	AEKEWELPPDSMDLNLIE--QAPLTVAEKQLQRELFVEWRRVSK--APEALFFVQFEKG-	85				
Qy	123	DQGMHCHVLIG-----GKDFSOAQGKWMRQLNVYMSRWLVLTACNVQLTPAERIKL	173				
Db	86	DSYFHLHLIVETGVKSMVGVRYVSQIKEK-----LVTRIYRGVEP-----	126				
Qy	174	RETAEDNEWTLTYKHQ--TKKDYTKCVLFGNMIAIYFLTKKKKISTSPRGGYFLSS	231				
Db	127	---QLPNPNFAVTKTRNGAGGNKVDDC-----YIPNYLLPK-----TQP-----	163				
Qy	232	DSGWK--TNF-----LKEGERHLVSKLYTDDMRPETVETTTTAQETKRGH-----	275				
Db	164	ELQWANTNMDOYISACLNLAERKELVAQHL-----THVSQTQEQNKQNPNSD	212				
Qy	276	---IQTKKEVSIKTLKELVHKRVTSPEDEMMQPOSYIEMMAQPGGNNLLKNTLEICTL	332				
Db	213	APVIRKTSARYMELVGLVDRGITSEKQWIQEQASYSIFNAASRSRQIKAAALDNASK	272				





QY 393 NAVLPHCPASTGKSIITAAQAVNGVCYNAAVNPPFNDCTKNLIWVEEAGNGQOV 452  
Db 328 NTIWLFGPATYTKNTIAEATAHAAVPPFGVCYNWTNENPPFNDCTKNLIWVEEAGKMTAKV 387  
QY 453 NQPKALCSGOTIRIDOKGKSKOIEPTPVIMTNENITVIRIGCEERPEHTQPIRDMLN 512  
Db 388 ESAKAILGSKVRVDQCKSSAQIDPTPVIIVTNTNCAVIDGNSITTFEHOQLQDMFK 447  
QY 513 IHTHLLPGDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWG--KVPDWE--NWAEPKV 569  
Db 448 FELTKRLEHDFGKVTQKQVDFRWDSDHVTETVTFYVRKGARKRPADNDADISEPKR 507  
QY 570 PTINLLGSARSPTTPKSTPLSQNYA 596  
Db 508 ACP-----SVAQPSSTAEPV--DYA 527  
RESULT 4  
US-09-807-802A-7  
; Sequence 7, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNVN.031USA  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US/09/807,802A  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 546  
; TYPE: PRT  
; ORGANISM: AAV-1  
US-09-807-802A-7

Query Match 11.8%; Score 428.5; DB 4; Length 546;  
Best Local Similarity 26.9%; Pred. No. 2.3e-35;  
Matches 156; Conservative 63; Mismatches 211; Indels 151; Gaps 20;  
QY 63 AETTDQSEDMEWE-----TTVDMTKKQVFIQDLSLVKCLFEVLNTKNIFPGDVNW 114  
Db 31 AEKELPPSDMDLNLIEQAPLTVAEKLDQDFLVQWRRVSKA-----PEALF 78  
QY 115 FVQHEGKQDQWCHVLIG-----GKDFSAQAGKWMR---RQNVVWSRWLVATCN 162  
Db 79 FVQFEGESY-FHLHLIVETTVGKSMVLGRFLSQIRDLVQTIYRGIEPTLPNFWAVT-- 135  
QY 163 VQITPAERIKLRIEADNEWVTLTYKHQTKDYKCVLFGNMIAYYFLTKKISTSP 222  
Db 136 -----KTRNGAGGKNKVDECY-----IPNYLLPK-----TOP- 163  
QY 223 RDGGYFLSSDSGWK--TNF-----LKEGERHLSKLYTDDMRPETVTTVTTAQETK 272  
Db 164 -----ELQAWNTMEEYISACLNLAEKRLVAQHL-----THVSQTQEQN 203  
QY 273 RGRIOTKKEYSI---KTTLKE-----LVHKRVTSPEDMMMQPDSDYEMMAQPGGNNLL 323  
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKQWIOEQDQASVIFSNASNSRSQI 263  
QY 324 KNTLEICTLTARTKTAFLILKEAETSKLTNFSLPDTRCFRIF-----A 368  
Db 264 KAALDNAGKIMALTAKSAPDVLVGPAPPA-----DIKTNRIYRILELNGYEPAYAGSV 315  
QY 369 FHGWNVYKVAICCVLNROGGKRNALPHGPASTGKSIITAAQAVNGVCYNAAVN 428  
Db 316 FLGW-----AQKRFGRKNTIWLFGPATYTKNTIAEATAHAAVPPYGCNVTNEN 363

QY 429 FPENDCTKNLIWVEEAGNGQOVNOKALCSGOTIRIDOKGKSKOIEPTPVIMTNEN 488  
Db 364 FPENDCTKNLIWVEEAGKMTAKYVESAKAILGSKVRVDQCKSSAQIDPTPVIIVTNTN 423  
QY 489 ITVIRIGCEERPEHTQPIRDMLNLIHTHLLPGDFGLVDKNWPMICAWLVKNGYQSTMA 548  
Db 424 MCAVIDGNSITTFEHOQLQDMFKFELTKRLEHDFGKVTQKQVDFRWDSDHVTETVTFYVRKGARKRPADNDADISEPKR 483  
QY 549 SYCAKWG-----KVPDWE--NWAEPKVPTTPKSTPLSQNYA 585  
Db 484 FVYVRKGARKRPADNDADK-SEPKRAC-----SVADPSTS 518  
RESULT 5  
US-09-807-802A-2  
; Sequence 2, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNVN.031USA  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US/09/807,802A  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: AAV-1  
US-09-807-802A-2

Query Match 11.8%; Score 428.5; DB 4; Length 623;  
Best Local Similarity 26.9%; Pred. No. 2.9e-35;  
Matches 156; Conservative 63; Mismatches 211; Indels 151; Gaps 20;  
QY 63 AETTDQSEDMEWE-----TTVDMTKKQVFIQDLSLVKCLFEVLNTKNIFPGDVNW 114  
Db 31 AEKELPPSDMDLNLIEQAPLTVAEKLDQDFLVQWRRVSKA-----PEALF 78  
QY 115 FVQHEGKQDQWCHVLIG-----GKDFSAQAGKWMR---RQNVVWSRWLVATCN 162  
Db 79 FVQFEGESY-FHLHLIVETTVGKSMVLGRFLSQIRDLVQTIYRGIEPTLPNFWAVT-- 135  
QY 163 VQITPAERIKLRIEADNEWVTLTYKHQTKDYKCVLFGNMIAYYFLTKKISTSP 222  
Db 136 -----KTRNGAGGKNKVDECY-----IPNYLLPK-----TOP- 163  
QY 223 RDGGYFLSSDSGWK--TNF-----LKEGERHLSKLYTDDMRPETVTTVTTAQETK 272  
Db 164 -----ELQAWNTMEEYISACLNLAEKRLVAQHL-----THVSQTQEQN 203  
QY 273 RGRIOTKKEYSI---KTTLKE-----LVHKRVTSPEDMMMQPDSDYEMMAQPGGNNLL 323  
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKQWIOEQDQASVIFSNASNSRSQI 263  
QY 324 KNTLEICTLTARTKTAFLILKEAETSKLTNFSLPDTRCFRIF-----A 368  
Db 264 KAALDNAGKIMALTAKSAPDVLVGPAPPA-----DIKTNRIYRILELNGYEPAYAGSV 315  
QY 369 FHGWNVYKVAICCVLNROGGKRNALPHGPASTGKSIITAAQAVNGVCYNAAVN 428  
Db 316 FLGW-----AQKRFGRKNTIWLFGPATYTKNTIAEATAHAAVPPYGCNVTNEN 363  
QY 429 FPENDCTKNLIWVEEAGNGQOVNOKALCSGOTIRIDOKGKSKOIEPTPVIMTNEN 488  
Db 364 FPENDCTKNLIWVEEAGKMTAKYVESAKAILGSKVRVDQCKSSAQIDPTPVIIVTNTN 423

Qy	489	ITVVRICCEERPEHTQPIRDRMLNIHLTHHLPGDFGLVDKNWPMLCAMLVKNQYOSTWA	548
Db	424	MCIVIDGNSSTTFEHQQPLODRMFKFELTRLEHDFGKVTQKEYKFFRWAQDHVTEVAHE	483
Qy	549	SYCAKWG- ---KVPDSENWAEKVPPTPINLLGSARSPTT	585
Db	484	FYRKGGANKRPAPDDADK-SEPKRACP-----SVADPSTS	518
RESULT 6			
US-09-807-802A-5			
; Sequence 5, Application US/09807802A			
; Patent No. 6759237			
; GENERAL INFORMATION:			
; APPLICANT: Willson, James M.			
; APPLICANT: Xiao, Weidong			
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,...			
; TITLE OF INVENTION: Vectors and Host Cells Containing Same			
; FILE REFERENCE: GNPVN.031USA			
; CURRENT APPLICATION NUMBER: US/09/807,802A			
; PRIOR FILING DATE: 2002-02-21			
; PRIOR APPLICATION NUMBER: US 60/107,114			
; PRIOR FILING DATE: 1998-11-05			
; PRIOR APPLICATION NUMBER: PCT/US99/25694			
; PRIOR FILING DATE: 1999-11-02			
; NUMBER OF SEQ ID NOS: 20			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 5			
; LENGTH: 623			
; TYPE: PRT			
; ORGANISM: AAV-1			
US-09-807-802A-5			
Query Match 11.8%; Score 428.5; DB 4; Length 623;			
Best Local Similarity 26.9%; Pred. No. 2.9e-35;			
Matches 156; Conservative 63; Mismatches 211; Indels 151; Gaps 20:			
Qy	63	AETTWOSEDMWE-----TTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNW	114
Db	31	AEKEWELPPDSMDLNLIEOAPLTVAEKLRDPLVQMRVSKA-----PEALF	78
Qy	115	FVQHEWKDQGWCHVLIG-----GKDFSOAGKKWR---ROLNVVSRWLVTACN	162
Db	79	FVQPEKESY-FHLHLIVETGVKSVLGRFSLQIRDKLVQIYRGIEFTLNNFAVT--	135
Qy	163	VQLTPAERIKLRBAEDNEWVTLITYKHQTKDYTKCVLFGNMIAYYELTKKISTGPP	222
Db	136	-----KTRNGAGGKNVDECY-----IPNYLLPK---TQP-	163
Qy	223	RDGGYFLSSDSGHW-KTNF-----LKEGERHLVSKLYTDDMRPETVETVTTAQTETK	272
Db	164	-----ELQAWNTNMEEYISACLNLAEKRLVAQHL-----THVSQTQEQN	203
Qy	273	RGRIOTKKEYSI---KTLIKE-----LVHKRVTPSPEDNMWMPDSYIEMMAQPGENLL	323
Db	204	KEWLNPNSDAPVTRSKTSARYMELVGLVDRGITSEKQWIQEDQASYISFNAASRSQI	263
Qy	324	KNTLEICTLTLARTKTAFDLILEKAFTSKLTNFSLPDTRCIF-----A	368
Db	264	KAALDNAGKIMALTGAPDYLVPAPPA-----DIKTNRIYRILELNGVEPAYAGSV	315
Qy	369	FHGWNVYKVCATCCVLNROGGKRNVLPHGPASTCKSIQAQIAQAVGNCVCYNAVN	428
Db	316	FLGW-----AQRFRGKNTIWLFGPATTTGKTNIAEIAHAHAFPFYGCVMNTNEN	363
Qy	429	FPFNDCTNKNLIWVEBAGNFGQOVNFKAICSQOTIRIDQKGGSKQIETPTVIMTTNEN	488
Db	364	FPENDCVDKMWIWESEGKMTAKVVEKAKILGSKVRVDQCKSSAQIDPTPIVTSNTN	423
Qy	489	ITVVRICCEERPEHTQPIRDRMLNIHLTHHLPGDFGLVDKNWPMLCAMLVKNQYOSTWA	548
Db	424	MCIVIDGNSSTTFEHQQPLODRMFKFELTRLEHDFGKVTQKEYKFFRWAQDHVTEVAHE	483

```

Qy 549 SYCAKWG-----KVPDWSENWAEPKVPTPINLLGARSPTT 585
Db 484 FYVRKGGANKRPAPDDADK-SEPKRACP-----SVADEPSTS 518

RESULT 7
US-09-532-594B-8
; Sequence 8, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532.594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 Rep protein 40
US-09-532-594B-8

Query Match 11.7%; Score 425.5; DB 4; Length 312;
Best Local Similarity 35.3%; Pred. No. 1.7e-35;
Matches 109; Conservative 45; Mismatches 140; Indels 15; Gaps 5

Qy 291 LVHKRVTSPEBMMMQPDSYIEMMAQGGENLLKNVTLEICTLTJLARTKTAFDLILEKAE 350
Db 7 LVDRGITSEKQWIGEDQASYSIFSAANSRSQIKAAALDNASKIMSLTKTAPDYLVGQNPP 66

Qy 351 SKLTNFSLPDTRTCRIFAFHGWNVYKVCHACCVLINRQGGKRNAVLPHGPASTGKSIQAQ 410
Db 67 EDISS-----NRIYRIEMNGYDPOYASVFLGWAQKAFGRKNTIWLFGPATGKNTIAE 121

Qy 411 AIAQAVNGVCYNAANYNFFPNDCTKNLWVEEAGNFGQOVNOFKAICSGQTIKIDQKG 470
Db 122 AIAHAVPFYGCWNVTNENFFPNDCCVDKMWIWEERGKMTAKVYESAKAILGSKVRVDQKC 181

Qy 471 KGSQIPEPTVIMTNNENITVVRIGCSERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE 530
Db 182 KSSAQIDPTFVIVTSNTNMCVIDGNSSTTFEHQQPFDQRMFKELTKRLEHDFGKVTQOE 241

Qy 531 WPMICAWLKNVGYSTMASYCAKWG--KVPDWS--NWAEPKVPTPINLLGARSPTTTPK 587
Db 242 VKDFFRWASDHVIEVTHFYVRKGGARKRPAPNDADISEPKRACP-----SVAQPTSDA 296

Qy 588 STPLSQNYA 596
Db 297 EAPV--DYA 303

RESULT 8
US-09-532-594B-9
; Sequence 9, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532.594B
; CURRENT FILING DATE: 2000-03-22

```

```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.6%; Score 425.5; DB 4; Length 399;
Best Local Similarity 34.7%; Pred. No. 2.7e-35;
Matches 109; Conservative 45; Mismatches 140; Indels 15; Gaps 5;

Qy 291 LVHKRVTSPEDMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEDQASVISFNAASNSRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66

Qy 351 SKLTNFSLPDTRTCRIF-----AFHGNVYKVKCHAIACCVLNROGGKENV 395
Db 67 A-----DIKTNRIYRILELNGYEPAYAGSVFLGM-----AKQRFGRKNTI 106

Qy 396 LFHGPASTGKSIIAQAIQAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQQVNOF 455
Db 107 WLFPGATTGKTNTAEIAHAHVPYGCNVNTNENFPFNDCCVCKWVIWEEGKMTAKVESA 166

Qy 456 KAICSGQTIRIDQKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDMNLNHL 515
Db 167 KAILGSKVRVDQCKSSAQIDPTPIVITSNTNCAVIDGNSITTFEHQQLQDRMFKFEL 226

Qy 516 THHLPDGLVDKNEPWCAMLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVP 571
Db 227 TRLEHDFGKVTQKVEKFFRWAQDHVTEVAHBFYVRKGGANKRPAPDDADK-SEPKRAC 285

Qy 572 PINLLGSARSPPFTT 585
Db 286 P-----SVADPSTS 294

RESULT 10
US-09-807-802A-9
; Sequence 9, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-9

Query Match      11.6%; Score 420.5; DB 4; Length 399;
Best Local Similarity 34.7%; Pred. No. 8.9e-35;
Matches 109; Conservative 37; Mismatches 123; Indels 45; Gaps 6;

Qy 291 LVHKRVTSPEDMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEDQASVISFNAASNSRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66

Qy 351 SKLTNFSLPDTRTCRIF-----AFHGNVYKVKCHAIACCVLNROGGKENV 395
Db 67 A-----DIKTNRIYRILELNGYEPAYAGSVFLGM-----AKQRFGRKNTI 106

Qy 396 LFHGPASTGKSIIAQAIQAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQQVNOF 455
Db 107 WLFPGATTGKTNTAEIAHAHVPYGCNVNTNENFPFNDCCVCKWVIWEEGKMTAKVESA 166

Qy 456 KAICSGQTIRIDQKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDMNLNHL 515
Db 167 KAILGSKVRVDQCKSSAQIDPTPIVITSNTNCAVIDGNSITTFEHQQLQDRMFKFEL 226

Qy 516 THHLPDGLVDKNEPWCAMLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVP 571
Db 227 TRLEHDFGKVTQKVEKFFRWAQDHVTEVAHBFYVRKGGANKRPAPDDADK-SEPKRAC 285

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.7%; Score 425.5; DB 4; Length 399;
Best Local Similarity 35.3%; Pred. No. 2.7e-35;
Matches 109; Conservative 45; Mismatches 140; Indels 15; Gaps 5;

Qy 291 LVHKRVTSPEDMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEDQASVISFNAASNSRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66

Qy 351 SKLTNFSLPDTRTCRIFAFHGNVYKVKCHAIACCVLNROGGKENVAFHGPASTGKSIIAQ 410
Db 67 EDLSS-----NRIYRILELNGYEPQYASVFLGWAQKFKGRNTIWLFGATTGKTNTAE 121

Qy 411 AIAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQQVNOFKAICSGQTIRIDQK 470
Db 122 AIAHAVPFYGCNVNTNENFPFNDCCVCKWVIWEEGKMTAKVESA KAILGSKVRVDQK 181

Qy 471 KGSQKTEPTVIMTNTNENITVVRIGCEERPEHTQPIRDMNLNHLTHHLPDGLVDKNE 530
Db 182 KSSAQIDPTPIVITSNTNCAVIDGNSITTFEHQQLQDRMFKFELTKRLEHDFGKVTQK 241

Qy 531 WPMICAWLVKNGYQSTMASYCAKMG--KVPDWSNNAEPKVPINLLGSARSPPFTPK 587
Db 242 VKDFFRWASHVTEVTHEFVYVRKGGARKRPAPDNADISEPKRACP-----SVAQPS 296

Qy 588 STPLSQNYA 596
Db 297 EAPV--DYA 303

RESULT 9
US-09-807-802A-11
; Sequence 11, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 322
; TYPE: PRT
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (943)..(944)
; OTHER INFORMATION: minor splice site
US-09-807-802A-11

Query Match      11.6%; Score 420.5; DB 4; Length 322;
Best Local Similarity 34.7%; Pred. No. 6e-35;
Matches 109; Conservative 37; Mismatches 123; Indels 45; Gaps 6;
```

QY 572 PINLGSARSPTT 585  
DB 286 P-----SVADPSTS 294

## RESULT 11

US-09-171-461-3  
; Sequence 3, Application US/09171461  
; Patent No. 6335016  
; GENERAL INFORMATION:

; APPLICANT: Baker, Adam  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Chiocca, Susanna  
; APPLICANT: Kurzbauer, Robert  
; APPLICANT: Schaffner, Gotthold  
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus  
; FILE REFERENCE: 0652.1800000  
; CURRENT APPLICATION NUMBER: US/09/171.461  
; CURRENT FILING DATE: 1999-01-12  
; EARLIER APPLICATION NUMBER: PCT/EP97/01944  
; EARLIER FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 276

; TYPE: PRT

; ORGANISM: CELO Virus

; FEATURE:

; OTHER INFORMATION: Position: 1999..2829 /note=ORF2

US-09-171-461-3

Query Match 3.8%; Score 139; DB 3; Length 276;  
Best Local Similarity 23.7%; Pred. No. 1.4e-05;  
Matches 63; Conservative 41; Mismatches 134; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEDEMMMQPDSY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343  
DB 10 TLVHALIDRGVVSREQWQVDPAQYQFYHRSKQGFK--VRHILRDVIRHMCWSTRLDDY 67  
QY 344 ILEKAETSKLTNFSLPD-----TTCRIFAFHGNVYKVCCHAICCVLNROGGKKNVLFHG 399  
DB 68 MSGASTPS-----PDDVLRNPLYQLLLCNGYNPAVVGTLIRWAGHQ--SNRNTVWIRG 119  
QY 400 PASTGKSIIAQAIAQAVGNVGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAIC 459  
DB 120 TPMSGAPYLAQAIAYCSPVSGVDWKNKSNPFEGCPDLSLVFWWDGGYVYDCCVGLVKQVF 179  
QY 460 SGQTRIDQKG----KGSQKIEPTPVIMTNNENITVVRIGCEE--RPEHTQPIRDRMLNTH 514  
DB 180 RGEHVILPPEGLRGNPCSELFTPTVLMYSQADICMTRLRSGLSABHAGVGLRDCWYLIR 239  
QY 515 LTHHLPGDF---GLVDKNEWPMICAW 537  
DB 240 LTE-----DFDCAGGISCADVQKQFVAV 261

## RESULT 12

US-09-970-711-3  
; Sequence 3, Application US/09970711  
; Patent No. 6773709  
; GENERAL INFORMATION:

; APPLICANT: Baker, Adam  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Chiocca, Susanna  
; APPLICANT: Kurzbauer, Robert  
; APPLICANT: Schaffner, Gotthold  
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus  
; FILE REFERENCE: 0652.1800001  
; CURRENT APPLICATION NUMBER: US/09/970.711  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/171.461  
; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: PCT/EP97/01944  
; PRIOR FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3

; LENGTH: 276

; TYPE: PRT

; ORGANISM: CELO Virus

; FEATURE:

; OTHER INFORMATION: Position: 1999..2829 /note=ORF2

US-09-970-711-3

Query Match 3.8%; Score 139; DB 4; Length 276;  
Best Local Similarity 23.7%; Pred. No. 1.4e-05;  
Matches 63; Conservative 41; Mismatches 134; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEDEMMMQPDSY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343  
DB 10 TLVHALIDRGVVSREQWQVDPAQYQFYHRSKQGFK--VRHILRDVIRHMCWSTRLDDY 67  
QY 344 ILEKAETSKLTNFSLPD-----TTCRIFAFHGNVYKVCCHAICCVLNROGGKKNVLFHG 399  
DB 68 MSGASTPS-----PDDVLRNPLYQLLLCNGYNPAVVGTLIRWAGHQ--SNRNTVWIRG 119  
QY 400 PASTGKSIIAQAIAQAVGNVGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAIC 459  
DB 120 TPMSGAPYLAQAIAYCSPVSGVDWKNKSNPFEGCPDLSLVFWWDGGYVYDCCVGLVKQVF 179  
QY 460 SGQTRIDQKG----KGSQKIEPTPVIMTNNENITVVRIGCEE--RPEHTQPIRDRMLNTH 514  
DB 180 RGEHVILPPEGLRGNPCSELFTPTVLMYSQADICMTRLRSGLSABHAGVGLRDCWYLIR 239  
QY 515 LTHHLPGDF---GLVDKNEWPMICAW 537  
DB 240 LTE-----DFDCAGGISCADVQKQFVAV 261

## RESULT 13

US-09-300-909-17  
; Sequence 17, Application US/09300909  
; Patent No. 6306580  
; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING

; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR

; NUMBER OF SEQUENCES: 27

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/300,909

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/083,942

; FILING DATE: 01-MAY-1998

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 647 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-300-909-17

Query Match 3.4%; Score 124.5; DB 3; Length 647;  
Best Local Similarity 19.7%; Pred. No. 0.0021;  
Matches 110; Conservative 78; Mismatches 172; Indels 197; Gaps 26;

QY 39 GKDIGNSYKKELODELK-SLQRAETTWQSE-----DMWEETT-VDEMKKQVFIIDS 92  
DB 100 GKVGGRNT-RGTLOEISLNVSTQATQTVSYSPDSGYGNMEVETAEVEVT----- 149

QY 93 LVKKCLFEVLNTKNIIPFGDNNVQVHEMGKQDQGHCHVLIGGKDFSOAQCKWRRRLNRY 152  
Db 150 -----VATNTNGDAEGHGGSVRBE-----CSSVDSASIDSENQDPK----- 185  
QY 153 WSRWLVACNVQVTPAERIKLRIADNEDNEWVTLTYKHKT-----KKDYTKC 200  
Db 186 -----SPTAQIKLL-LQSNKKAAMLT-OFKETYGLSFTDLVTRTFKSOKTTC 230  
QY 201 VLFGNMAYVFLTKKISTSPRPGDGYFLSSDSGWKTNFLKEGERHLVSK--LYTODMRP 258  
Db 231 T----DWVAALFGVHPITIA-----EGFKTLINKYALYTH----- 260  
QY 259 ETVETVTTTAAETKRGRIOT-----KKEVSIKTKTLKELVHKRVTSPEDEMMQPDYSY 310  
Db 261 -----IQSLDTKQGVLLILMLIRYTCGRNVRTVGKGLSTLLH-----VPESCMLLEPPK 309  
QY 311 IEMMA-----OPGENLLKNTLE-----ICTLTAR---TKTAFDL--ILEKAETSKLT-- 354  
Db 310 RSPVAALYVTRTGISNISVVGTGPWIRQLTVIQHGIDDSVDFLSDMVQWAFDNEYTDE 369  
QY 355 -----NFSLPDTRTCRIFAFHGNW-----YVKVCHAIC-----CVL 385  
Db 370 SDIAFYVAMLADCSNAAFLKNSCOAKYVYKDCATWCKHYKRAQKQMSMSOWIKERCCK 429  
QY 386 NRQGG-----KRNALVPHGPASTGKSIIAQIAQ 414  
Db 430 CDEGGDWRPVQFLRYQGIIFISFLCALKEFLKGTGPKKNCIVYGPANTGKSHFCMSLMH 489  
QY 415 AV-GNVGCYNAANVNPFDCTNKNLIWVEAGN--FGQVNVQFKAICSGQTRIRIDOKGK 471  
Db 490 FLOQTIVSYVNSTSHFWLEPLADAKLMLDDATGTCTWSYFNDYMRNALDGYAISLDRKYK 549  
QY 472 GSKQIEPTVPVMTNEN 488  
Db 550 SLLQMKCPPLITSNTN 566

## RESULT 14

US-09-134-001C-2897  
; Sequence 2897, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 2897  
; LENGTH: 564  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2897

Query Match 3.1%; Score 114; DB 3; Length 564;  
Best Local Similarity 21.4%; Pred. No. 0.021;  
Matches 81; Conservative 63; Mismatches 128; Indels 106; Gaps 21;  
QY 188 YKHQTKDYTKCVLFGNMTAYFLTKKISTSPRGGGYFLSSDSGW-KTNFLKEGE-- 244  
Db 87 YSYFGTKGNPNDVLMEMMIKVPHI--KRFNSE-----LFVPKANGWQKIN----GDEL 134  
QY 245 RHLVSKL---YTDMPRETVEVTTTAQETKRGRIQTKKEVSIKTKTLKELVHKRVTSPE 301  
Db 135 QGLISKMIQVLLVDYKPSL-----STLKNVVDGLQKSTDV 169  
QY 302 WMMMQPDSYIEMMAQPGENL-----LKNLTLEICTLT-LARTKTAFLILEKAB--- 349

Db 170 EBLVENEHVGIC-----GENMFDLNTFQVVKNSIDIFPKTRLNLSLSTNDVITDKIPPPYF 224  
QY 350 ---TSKLTNFSPLPDRTRTCRIFAFHGNVYKVAICCVLNRQGGKKNVAVLFGHPASTGKS 406  
Db 225 KQYMLQLANYD---DDLQYFLFQ-----HT-AVLLTADTKYRGLLYGAKNGKS 271  
QY 407 IIAQIAQAVGNVGCYNAANVNPFD-----CTNKNLIWVEAGNFGQO---VNOQ 454  
Db 272 VVIELVKSPF---YSKDIVSKPLNELGFRFDKESLIDKSLMASHEIGOSKIQEIIVND 326  
QY 455 FKAICSGQTRIDOKGSKQJE---PTPVMTNENITVVRIGCERPEHTOPIRDRML 511  
Db 327 FKLLSVESMHVRDKGK--TQVEVILDLKLIFSTNAILNF-----PPEHAKALERRIN 377  
QY 512 NIHLTHL-PGDFGLVVK 528  
Db 378 IIPCEYVVEKADTSLDK 395

## RESULT 15

US-09-107-532A-6086  
; Sequence 6086, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6086:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (B) LOCATION 1...589  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6086:  
US-09-107-532A-6086

Query Match 3.1%; Score 113; DB 4; Length 589;  
Best Local Similarity 19.8%; Pred. No. 0.029;

Matches	93; Conservative	61; Mismatches	152; Indels	164; Gaps	22;
Qy	162 NVQLTPAEIRIKLRETAEDNEWWTLLT-YKHQTKDYTKCVLFGNMIAYFLTKKISTS	220			
Db	142 NIKLRNATTKAIRHYLDNDPDIETPYLGKSTPEG-----ARDYLVPSRV---	187			
Qy	221 PPRGGYFLS---SDSGHKTNFKEG--ERHLVSKLYTD-----DMRPE-----TVETVT	266			
Db	198 ---HAGHYFALPQSPQLFKQLMAGDFRYQIVRCFRDEDLRGDRQPEFTQVDIETTF	244			
Qy	267 TAQETKRGRIQTKKEVSITKTLKELVHKRVTSPEDEMMMQPDSYIEMMAQPGGENLLKNT	326			
Db	245 TAEE-----IQTYTEGLIAKVWKEVRGIEVTLPPRM-----TYDEAMARYGSD-----	289			
Qy	327 LEICTLTARTKTAFDL-ILEKAETSKLTNFSLEPDTTRCTRIFAHHGWNVYKVCHAI CCVL	385			
Db	289 -----KPDTRFDMELIDLSDTVKEVEF-----KVFO-----MA	316			
Qy	386 NRQGGKRNAVLPHGPASTGKSIIAQATAQAVGNVGCYNAAANVPFNDCTNKNLIW--VE	443			
Db	317 LENGGVKALNAKGAADRYSRKMDQLGQYVGQFGA-----KGLAWLKVE	361			
Qy	444 EAGNFGQOVNOFKAICSGQIRIDQKGSQIEPTVIM--TTNENITVVRIGCEERPE	501			
Db	362 EDGLKG-PIAKPMGEATEAIK-----ATDAKPGDILMFGADKSEIVAAALGA-----	408			
Qy	502 HTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVKNGYQSTMASYCAKWKGVDPWS	561			
Db	409 ----IRTR-----LGKELGLIDESKFNFL-----WV	430			
Qy	562 ENWAEPKVPTPINLLGARSPTTPKSTPLSONYALTPLASDLEDLALEP	611			
Db	431 TDWPFQFESEEGRYVSAHHPTMPKE-----EDIPLLAEDP	467			

Search completed: January 22, 2005, 03:43:22  
 Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:41:01 ; Search time 90.6667 Seconds  
(without alignments)  
2677.792 Million cell updates/sec

Title: US-10-069-056-14  
Perfect score: 3637  
Sequence: 1 MAGNAYDEVILGATNWLKKE.....RACFAGAEPLKXOFSEPLNLD 672

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	459.5	12.6	626	9	US-09-792-630-29
2	459.5	12.6	626	10	US-09-953-351-29
3	459.5	12.6	626	13	US-10-080-376-29
4	459.5	12.6	626	14	US-10-082-671-35
5	459.5	12.6	626	14	US-10-097-100-29
6	459.5	12.6	626	14	US-10-023-208-29
7	459	12.6	627	9	US-09-792-630-25
8	459	12.6	627	10	US-09-953-351-25
9	459	12.6	627	13	US-10-080-376-25
10	459	12.6	627	14	US-10-082-671-31
11	459	12.6	627	14	US-10-097-100-25
12	459	12.6	627	14	US-10-023-208-25
13	453.5	12.5	610	9	US-09-792-630-21

14	453.5	12.5	610	10	US-09-953-351-21	Sequence 21, Appl
15	453.5	12.5	610	13	US-10-080-376-21	Sequence 21, Appl
16	453.5	12.5	610	14	US-10-082-671-27	Sequence 27, Appl
17	453.5	12.5	610	14	US-10-097-100-21	Sequence 21, Appl
18	453.5	12.5	610	14	US-10-023-208-21	Sequence 21, Appl
19	453.5	12.5	610	14	US-10-375-192-7	Sequence 7, Appl
20	440.5	12.1	625	14	US-10-423-704A-3	Sequence 3, Appl
21	440.5	12.1	627	9	US-09-792-630-27	Sequence 27, Appl
22	440.5	12.1	627	10	US-09-953-351-27	Sequence 27, Appl
23	440.5	12.1	627	13	US-10-080-376-27	Sequence 27, Appl
24	440.5	12.1	627	14	US-10-082-671-33	Sequence 33, Appl
25	440.5	12.1	627	14	US-10-097-100-27	Sequence 27, Appl
26	440.5	12.1	627	14	US-10-023-208-27	Sequence 27, Appl
27	432.5	11.9	537	10	US-09-254-747-10	Sequence 10, Appl
28	432.5	11.9	623	9	US-09-792-630-5	Sequence 5, Appl
29	432.5	11.9	623	10	US-09-953-351-5	Sequence 5, Appl
30	432.5	11.9	623	10	US-09-254-747-2	Sequence 2, Appl
31	432.5	11.9	623	10	US-09-254-747-11	Sequence 11, Appl
32	432.5	11.9	623	13	US-10-080-376-5	Sequence 5, Appl
33	432.5	11.9	623	14	US-10-082-671-11	Sequence 11, Appl
34	432.5	11.9	623	14	US-10-097-100-5	Sequence 5, Appl
35	432.5	11.9	623	14	US-10-023-208-5	Sequence 5, Appl
36	432.5	11.9	623	14	US-10-375-192-5	Sequence 5, Appl
37	431.5	11.9	623	9	US-09-792-630-13	Sequence 13, Appl
38	431.5	11.9	623	10	US-09-953-351-13	Sequence 13, Appl
39	431.5	11.9	623	13	US-10-080-376-13	Sequence 13, Appl
40	431.5	11.9	623	14	US-10-082-671-19	Sequence 19, Appl
41	431.5	11.9	623	14	US-10-097-100-13	Sequence 13, Appl
42	431.5	11.9	623	14	US-10-023-208-13	Sequence 13, Appl
43	431.5	11.9	623	14	US-10-375-192-2	Sequence 2, Appl
44	428.5	11.8	546	15	US-10-696-261-7	Sequence 7, Appl
45	428.5	11.8	546	15	US-10-696-282-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-09-792-630-29  
; Sequence 29, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahiyat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-09-792-630-29

Query Match 12.6%; Score 459.5; DB 9; Length 626;  
Best Local Similarity 25.3%; Pred. No. 6.3e-31;  
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

QY	75	WETT-----VDEMTKKQVFIFDSLVKCLFEVLTKNIF-----PGDVNWNVQHEWGK	122
Db	45	WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQWNNQFNQDETDFDFOLEEGS	93
QY	123	DQGWCHVLIGKDFSOAQGWRRQLNVYWSRWLTACNVQLTPAERIK---LRETAED	179
Db	94	EY-IHLHVCPCGECRSFVLGRY-----MSQRKDSLRLDYEG	129
QY	180	N-----EWTLTYKHQTKDYTKCVLFGNNMIAYVFLTKKISTSPPRDGGYFLSSDSG	234
Db	130	KQVKIPDFWSIYTKKGGQNKVTAA-----YILHLVLPKKQ-----PELOWAFTNPLF	179
QY	235	WKTNFLKEGERHLVSKLYTDDMRPETVTTVTTAQTGKRGIQTKKEVSIKTTKLVLHK	294

Db 180 TAAALCLQKRELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKVSNLVDWLIEM 235  
Qy 295 RVTSPEDEMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFADLILEKAETSKLT 354  
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295  
Qy 355 NFSLPDRTRCIRIFAFHGMNVYKCHAI CCVLNRQGGKRNALVPHGPASTGKSIIAQIAQ 414  
Db 296 K-----NRIYQILKLNYPQYVGVLCGWKREFNKRNAIWLYPATTKTNIAEIAH 350  
Qy 415 AVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGGSK 474  
Db 351 AVFPYGCNVNTNENFPFND CVDKMLIWMEEGKNTKRVESAKAILGGSRAVRVDKCKGSV 410  
Qy 475 QIEPTPVMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530  
Db 411 CIEPTPVITSTNDCMIVDGNSTTMEHRIPLERMFQIVLSHKGEGFKISKKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568  
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPKIWAAPTRELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDALEPWTSPNT 617  
Db 525 LRASPELFASVAPLP-----SSPDTSPKRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTATQNTGAGSKAC-QDQQLSPTWSEIEEDLRACFGAEPLKDPSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

## RESULT 2

US-09-953-351-29  
; Sequence 29, Application US/09953351  
; Publication No. US20030036643A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I  
; FILE REFERENCE: A-70814/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/953,351  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232,960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-09-953-351-29

Query Match 12.6%; Score 459.5; DB 10; Length 626;  
Best Local Similarity 25.3%; Pred. No. 6.3e-31;  
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;  
Qy 75 WETT-----VDEMTKKQVIFDPSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWNMEHVNLPMTLADKI-----KNIFIQRMNQFNQDETDFPFQLEEGS 93  
Qy 123 DQGWCHVLIGGKDFSAQAGKWRRLQNVVSRWLVTACNVQLTPAERIK---LREIAED 179  
Db 94 EY-IHLHAVCPGRCRPFVLGRY-----MSQIKDSILRDVYEG 129  
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYVFLTKKISTSPRGGYFLSDSG 234  
Db 130 KQVKIPDWFISITTKRGQGNKVTAA-----YILHYLIPKKQ-----PELOWAFTNPLF 179  
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVTTVTAQETKRGRIQTKKEVSIKTKLKVHVK 294

Db 180 TAAALCLQKRELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKVSNLVDWLIEM 235  
Qy 295 RVTSPEDEMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFADLILEKAETSKLT 354  
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295  
Qy 355 NFSLPDRTRCIRIFAFHGMNVYKCHAI CCVLNRQGGKRNALVPHGPASTGKSIIAQIAQ 414  
Db 296 K-----NRIYQILKLNYPQYVGVLCGWKREFNKRNAIWLYPATTKTNIAEIAH 350  
Qy 415 AVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIRIDQKGGSK 474  
Db 351 AVFPYGCNVNTNENFPFND CVDKMLIWMEEGKNTKRVESAKAILGGSRAVRVDKCKGSV 410  
Qy 475 QIEPTPVMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530  
Db 411 CIEPTPVITSTNDCMIVDGNSTTMEHRIPLERMFQIVLSHKGEGFKISKKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568  
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPKIWAAPTRELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDALEPWTSPNT 617  
Db 525 LRASPELFASVAPLP-----SSPDTSPKRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTATQNTGAGSKAC-QDQQLSPTWSEIEEDLRACFGAEPLKDPSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

## RESULT 3

US-10-080-376-29  
; Sequence 29, Application US/10080376  
; Publication No. US20020172968A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahiyat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/080,376  
; CURRENT FILING DATE: 2000-02-19  
; PRIOR APPLICATION NUMBER: US 09/792,630  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-10-080-376-29

Query Match 12.6%; Score 459.5; DB 13; Length 626;  
Best Local Similarity 25.3%; Pred. No. 6.3e-31;  
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;  
Qy 75 WETT-----VDEMTKKQVIFDPSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWNMEHVNLPMTLADKI-----KNIFIQRMNQFNQDETDFPFQLEEGS 93  
Qy 123 DQGWCHVLIGGKDFSAQAGKWRRLQNVVSRWLVTACNVQLTPAERIK---LREIAED 179  
Db 94 EY-IHLHAVCPGRCRPFVLGRY-----MSQIKDSILRDVYEG 129  
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYVFLTKKISTSPRGGYFLSDSG 234  
Db 130 KQVKIPDWFISITTKRGQGNKVTAA-----YILHYLIPKKQ-----PELOWAFTNPLF 179  
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVTTVTAQETKRGRIQTKKEVSIKTKLKVHVK 294  
Db 180 TAAALCLQKRELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKVSNLVDWLIEM 235  
Qy 295 RVTSPEDEMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFADLILEKAETSKLT 354





Db 236 GITSEKOWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295  
Qy 355 NFSLPDTRTCRIFAFPHGWNVYKVCCHAICCVLNROGGKRNALVPHGPASTGKSIIAQIAQ 414  
Db 296 K-----NRIYQILKLNYPQYVGVSLGWVKEFNKRNALWLYGPATTKTNIAEAIAH 350  
Qy 415 AVGNVCYNAANVPFNDCTNKNLWVEEAGNFQGVNQFKAICSGQTTIRIDQKGGSK 474  
Db 351 AVPFYGCNVMTNENFPFNDCTNKNLWVEEAGNFQGVNQFKAICSGQTTIRIDQKGGSK 410  
Qy 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530  
Db 411 CIEPTVILTSNTDMCMVDGNSSTTWEHRIPLEERFQIVLSHKLGNFGKISKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKMGKVPDWSN-----WAEPK----- 568  
Db 471 FKWANDNLVPVSEFKVPTNEQTKLE-----PVERANEPSEPPKINAPPTRELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLIEDLALFPWSTPNT 617  
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTAEONTGEGSKAC-QDQGLSPTWSEIEEDLRACFGABPLKDDFSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFNCLEKDDMNE 612

## RESULT 6

US-10-023-208-29  
; Sequence 29, Application US/10023208  
; Publication No. US20030124537A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Liu, Yuan-Ching  
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES  
; FILE REFERENCE: A-70174-1/RET/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/023,208  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/256,163  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 29  
; TYPE: PRT  
; ORGANISM: Muscovy duck parvovirus  
US-10-023-208-29

Query Match 12.6%; Score 459.5; DB 14; Length 626;  
Best Local Similarity 25.3%; Pred. No. 6.3e-31;  
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;  
Qy 75 WETT-----VDEMTKKQVFIFDLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWNHVNLPWTLADKI-----KNIFIQWQNFQNDTDFFFOLEGS 93  
Qy 123 DQGWCHVLIGKGFDSQAQGWRRQLNVYWSRWLVATCNVOLTPAERIK---LREIAD 179  
Db 94 EY-IHLHVCPCGCRSPVLGRY-----NSQIKDSILRDVYEG 129  
Qy 180 N-----EWVLLTYKHQTKDYTKCVLFGNMIAVYFLTKKISTSPPRDGGVFLSSDSG 234  
Db 130 KQVKIPDWFISITTKRGGQNKVTAA-----YILHVLIPKKQ-----PELQWAFNTMPLF 179  
Qy 235 WKTNFKEGRHLVSKLYTDDMRPETVTTVTAQETKRGRIQTKKEVSIKTKLKVHK 294  
Db 180 TAAALCQKQELLDAFQSEMNNAVQEDQASTAPL-----ISNRAKNYSNLVDWLIEM 235  
Qy 295 RVTSPEDDMMQPDSDYIEMMAQPGGENLKNLTLEICTLTARTKTAFDILILEKAETSKLT 354  
Db 236 GITSEKOWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295  
Qy 355 NFSLPDTRTCRIFAFPHGWNVYKVCCHAICCVLNROGGKRNALVPHGPASTGKSIIAQIAQ 414

Db 296 K-----NRIYQILKLNYPQYVGVSLGWVKEFNKRNALWLYGPATTKTNIAEAIAH 350  
Qy 415 AVGNVCYNAANVPFNDCTNKNLWVEEAGNFQGVNQFKAICSGQTTIRIDQKGGSK 474  
Db 351 AVPFYGCNVMTNENFPFNDCTNKNLWVEEAGNFQGVNQFKAICSGQTTIRIDQKGGSK 410  
Qy 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530  
Db 411 CIEPTVILTSNTDMCMVDGNSSTTWEHRIPLEERFQIVLSHKLGNFGKISKEVKEF 470  
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKMGKVPDWSN-----WAEPK----- 568  
Db 471 FKWANDNLVPVSEFKVPTNEQTKLE-----PVERANEPSEPPKINAPPTRELEEI 524  
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLIEDLALFPWSTPNT 617  
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576  
Qy 618 PVAGTAEONTGEGSKAC-QDQGLSPTWSEIEEDLRACFGABPLKDDFSE 667  
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFNCLEKDDMNE 612

## RESULT 7

US-09-792-630-25  
; Sequence 25, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahiyat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RET/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 25  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-09-792-630-25

Query Match 12.6%; Score 459; DB 9; Length 627;  
Best Local Similarity 26.0%; Pred. No. 6.9e-31;  
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;  
Qy 75 WETT-----VDEMTKKQVFIFDLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWNHVNLPWTLADKI-----KNIFIQWQNFQNDTDFFFOLEGS 93  
Qy 123 DQGWCHVLIGKGFDSQAQGWRRQLNVYWSRWLVATCNVOLTPAERIK---LREIAD 179  
Db 94 EY-IHLHCCI-----AGNVRSFVLGRYS-----QIKDSILRDVYEG 130  
Qy 180 N-----EWVLLTYKHQTKDYTKCVLFGNMIAVYFLTKKISTSPPRDGGVFLSSDSG 234  
Db 131 KQVKIPDWFISITTKRGGQNKVTAA-----YILHVLIPKKQ-----PELQWAFNTMPLF 180  
Qy 235 WKTNFKEGRHLVSKLYTDDMRPETVTTVTAQETKRGRIQTKKEVSIKTKLKVHK 294  
Db 181 TAAALCQKQELLDAFQSEMNNAVQEDQASTAPL-----ISNRAKNYSNLVDWLIEM 236  
Qy 295 RVTSPEDDMMQPDSDYIEMMAQPGGENLKNLTLEICTLTARTKTAFDILILEKAETSKLT 354  
Db 237 GITSEKOWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 296  
Qy 355 NFSLPDTRTCRIFAFPHGWNVYKVCCHAICCVLNROGGKRNALVPHGPASTGKSIIAQIAQ 414  
Db 297 K-----NRIYQILKLNYPQYVGVSLGWVKEFNKRNALWLYGPATTKTNIAEAIAH 351  
Qy 415 AVGNVCYNAANVPFNDCTNKNLWVEEAGNFQGVNQFKAICSGQTTIRIDQKGGSK 474

Db 352 AVPFYGCWNTNENFPFNDVCDKMLIWWBEGKMTNKVESAAILGGSAVRVDQCKGVS 411  
QY 475 QIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNE---- 530  
Db 412 CIEPTPVIIITSNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGEFGFKISKEVKEF 471  
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568  
Db 472 FKWANDNLVPPVSEFKVPTNEQTKLTE-----PVPERANEPEPPKIWAAPTREBLEEI 525  
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPWPSTPNT 617  
Db 526 LRASPELFASVAPLP-----SSPDTSPKXKTRGEYQVRCAMHSL-DNSMNVFECLEC 577  
QY 618 PVAGTAETQNTGAGSKAC-QDQLSPTWSEIBEDLRACFAGBAPLKXDFSE 667  
Db 578 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

## RESULT 8

US-09-953-351-25

; Sequence 25, Application US/09953351

; Publication No. US20030036643A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min

; APPLICANT: Melander, Christian

; APPLICANT: Liu, Hong-Xiang

; APPLICANT: Jin, Cheng He

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION

; FILE REFERENCE: A-70814/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/953,351

; PRIOR FILING DATE: 2001-09-14

; PRIOR APPLICATION NUMBER: US 60/232,960

; PRIOR FILING DATE: 2000-09-14

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Barbarie duck parvovirus

US-09-953-351-25

Query Match 12.6%; Score 459; DB 10; Length 627;

Best Local Similarity 26.0%; Pred. No. 6.9e-31;

Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVIFDLSLVKKCLFEVLNTKNIF-----PGDVNVFVQHEWGK 122  
Db 45 WEPTGIWNMEHVNLPMVTLADKI-----KNIFIQRMNQFNQDETDFFFLQLEEGS 93

QY 123 DQGWCHVLIGGKDFSOAQGKWRRLQNVYWSRWLVLTACNVQLTPAERIK---LREIAED 179  
Db 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVTLTYKHQTKQYTKCVLFGNMIAYFYLTKKKISTSPPRDGGYFLSSDSG 234  
Db 131 KQVKIPDWSITTKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 180

QY 235 WKTNFKEGEHLVSKLYTDDMRPETVTTTAQETKRGRIQTKKEVSIKTKLKVHVK 294  
Db 181 TAAALCLQKQELLDAPQESEMNAVQEDQAAPL-----ISNRAAKNYSNLDVLIEM 236

QY 295 RVTSPEDDMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354  
Db 237 GITSEKQWLTENKESYRSFOATSSNNRQVKAALENABEMLLTKTATDYLGKDPVLDT 296

QY 355 NFSLPDTRTCRIPAFHGWNVYKCHAI CCVLNRRQGGKRNVAFLPHGPASTGKSIIAQIAQ 414  
Db 297 K-----NRIYQLKLNYPQYVGVSLCGWVKREFNKNAILWLYGPATTKTNIAEALAH 351

QY 415 AVNGVGCYNAANVPFNDCTNKNLIWVEAGNFGQOVNQFKALICSGQTTIRIDOKGKSK 474  
Db 352 AVPFYGCWNTNENFPFNDVCDKMLIWWBEGKMTNKVESAAILGGSAVRVDQCKGVS 411

## RESULT 9

US-10-080-376-25

; Sequence 25, Application US/10080376

; Publication No. US20020172968A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min

; APPLICANT: Bahiyat, Basil I.

; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES

; FILE REFERENCE: A-70295-2/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/10/080,376

; CURRENT FILING DATE: 2000-02-19

; PRIOR APPLICATION NUMBER: US 09/792,630

; PRIOR FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 87

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Barbarie duck parvovirus

US-10-080-376-25

Query Match 12.6%; Score 459; DB 13; Length 627;

Best Local Similarity 26.0%; Pred. No. 6.9e-31;

Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVIFDLSLVKKCLFEVLNTKNIF-----PGDVNVFVQHEWGK 122  
Db 45 WEPTGIWNMEHVNLPMVTLADKI-----KNIFIQRMNQFNQDETDFFFLQLEEGS 93

QY 123 DQGWCHVLIGGKDFSOAQGKWRRLQNVYWSRWLVLTACNVQLTPAERIK---LREIAED 179  
Db 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVTLTYKHQTKQYTKCVLFGNMIAYFYLTKKKISTSPPRDGGYFLSSDSG 234  
Db 131 KQVKIPDWSITTKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 180

QY 235 WKTNFKEGEHLVSKLYTDDMRPETVTTTAQETKRGRIQTKKEVSIKTKLKVHVK 294  
Db 181 TAAALCLQKQELLDAPQESEMNAVQEDQAAPL-----ISNRAAKNYSNLDVLIEM 236

QY 295 RVTSPEDDMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354  
Db 237 GITSEKQWLTENKESYRSFOATSSNNRQVKAALENABEMLLTKTATDYLGKDPVLDT 296

QY 355 NFSLPDTRTCRIPAFHGWNVYKCHAI CCVLNRRQGGKRNVAFLPHGPASTGKSIIAQIAQ 414  
Db 297 K-----NRIYQLKLNYPQYVGVSLCGWVKREFNKNAILWLYGPATTKTNIAEALAH 351

QY 415 AVNGVGCYNAANVPFNDCTNKNLIWVEAGNFGQOVNQFKALICSGQTTIRIDOKGKSK 474  
Db 352 AVPFYGCWNTNENFPFNDVCDKMLIWWBEGKMTNKVESAAILGGSAVRVDQCKGVS 411

QY 475 QIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNE---- 530  
Db 412 CIEPTPVIIITSNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGEFGFKISKEVKEF 471

Db 412 CIEPTPVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEP 471  
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPWSEN-----WAEPK----- 568  
Db 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPPSEPKIWAAPTREELLEE 525  
Qy 569 -----VPTPINLGSARSPTTPKSTPLSONYALTPLASDLELALPEWSTPNT 617  
Db 526 LRASPELFASVAPLP-----SSPDTSPKPKKTRGEYQVRCAMHSL-DNSMNVFECLEC 577  
Qy 618 PVAGTATONTGEAGSKAC-ODGQLSPTWSEIEEDLRACFGABPLKPDFSE 667  
Db 578 ERANPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 10  
US-10-082-671-31  
; Sequence 31, Application US/10082671  
; Publication No. US20030049647A1  
; GENERAL INFORMATION:  
; APPLICANT: DAHIYAT, BASSIL  
; APPLICANT: LI, MIN  
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL  
; TITLE OF INVENTION: PROFILES  
; FILE REFERENCE: XEN/001  
; CURRENT APPLICATION NUMBER: US/10/082,671  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/270,781  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 31  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-10-082-671-31

Query Match 12.6%; Score 459; DB 14; Length 627;  
Best Local Similarity 26.0%; Pred. No. 6.9e-31;  
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;  
Qy 75 WETT-----VDGMTKKQVFIQFDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122  
Db 45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRMNQFNQDTEDFDFQLEEGS 93  
Qy 123 DQGWCHVLIGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTPAERIK---LREIAD 179  
Db 94 EY-IHLHCCI-----AQGNVRSFVLGRYMS-----QIKDSILRDVYEG 130  
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNNIAYVFLTKKISTSPPRDGGYFLSSDSG 234  
Db 131 KQVKIPDWFISITTKRGQGNKVTAA-----YILHVLIPKKQ-----PELQWAFNTMPLF 180  
Qy 235 WKTNFLKEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOTKKEVSIKTLKELVHK 294  
Db 181 TAAALCLQKROELLDAFOSEMNAVQEOQAATAPL-----ISNRAAKNYSNLVDWLIEM 236  
Qy 295 RVTSPEDMMQPDSEYEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354  
Db 237 GITSEKWLTKENKESYRSFQATSSNNRQVKALENARAEMLLTATDYLIGKDPVLDIT 296  
Qy 355 NFSLPDTRTCRIFAFGWNVVYVCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQ 414  
Db 297 K-----NRIYQILKLNYNPQYVGVSLCGWVKREFNKRNAIWLPGPATTKTNAIEAIAH 351  
Qy 415 AVGNVCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDOKGKSK 474  
Db 352 AVFYGCNVNTNENFPND CVDKMLIWEEGKMTNKVSESAKAILGGSVAVRDQCKGKV 411  
Qy 475 QIEPTPVIMTNTENITVVRIGCEERPEHTOPIDRMLNIHLTHLPDGFGLVDKNE---- 530  
Db 412 CIEPTPVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEP 471

Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPWSEN-----WAEPK----- 568  
Db 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPPSEPKIWAAPTREELLEE 525  
Qy 569 -----VPTPINLGSARSPTTPKSTPLSONYALTPLASDLELALPEWSTPNT 617  
Db 526 LRASPELFASVAPLP-----SSPDTSPKPKKTRGEYQVRCAMHSL-DNSMNVFECLEC 577  
Qy 618 PVAGTATONTGEAGSKAC-ODGQLSPTWSEIEEDLRACFGABPLKPDFSE 667  
Db 578 ERANPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 11  
US-10-097-100-25  
; Sequence 25, Application US/10097100  
; Publication No. US20030068649A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Melander, Christian  
; APPLICANT: Liu, Hong-Xiang  
; APPLICANT: Jin, Cheng He  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI  
; FILE REFERENCE: A-70814/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/097,100  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/953,351  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: US 60/232,960  
; PRIOR FILING DATE: 2000-09-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 25  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-10-097-100-25

Query Match 12.6%; Score 459; DB 14; Length 627;  
Best Local Similarity 26.0%; Pred. No. 6.9e-31;  
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;  
Qy 75 WETT-----VDGMTKKQVFIQFDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122  
Db 45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRMNQFNQDTEDFDFQLEEGS 93  
Qy 123 DQGWCHVLIGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTPAERIK---LREIAD 179  
Db 94 EY-IHLHCCI-----AQGNVRSFVLGRYMS-----QIKDSILRDVYEG 130  
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNNIAYVFLTKKISTSPPRDGGYFLSSDSG 234  
Db 131 KQVKIPDWFISITTKRGQGNKVTAA-----YILHVLIPKKQ-----PELQWAFNTMPLF 180  
Qy 235 WKTNFLKEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOTKKEVSIKTLKELVHK 294  
Db 181 TAAALCLQKROELLDAFOSEMNAVQEOQAATAPL-----ISNRAAKNYSNLVDWLIEM 236  
Qy 295 RVTSPEDMMQPDSEYEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354  
Db 237 GITSEKWLTKENKESYRSFQATSSNNRQVKALENARAEMLLTATDYLIGKDPVLDIT 296  
Qy 355 NFSLPDTRTCRIFAFGWNVVYVCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQ 414  
Db 297 K-----NRIYQILKLNYNPQYVGVSLCGWVKREFNKRNAIWLPGPATTKTNAIEAIAH 351  
Qy 415 AVGNVCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDOKGKSK 474  
Db 352 AVFYGCNVNTNENFPND CVDKMLIWEEGKMTNKVSESAKAILGGSVAVRDQCKGKV 411  
Qy 475 QIEPTPVIMTNTENITVVRIGCEERPEHTOPIDRMLNIHLTHLPDGFGLVDKNE---- 530  
Db 412 CIEPTPVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEP 471

QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568  
Db 472 FKWANDNLVPVSEFKVPTNEOTKLTE-----PVPERANEPSEPPKIWAPPTRELEEEI 525  
QY 569 -----VPTPINLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPWPSTPNT 617  
Db 526 LRASPELFASVAPLP-----SSPDTSPKPKTRGEYQVRCAMHSL-DNSMNVFECLEC 577  
QY 618 PVAGTAPTONTGAGSKAC-ODQOLSPTWSEIEEDLRACFGAPLKKDFSE 667  
Db 578 ERANFPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 12  
US-10-023-208-25  
; Sequence 25, Application US/10023208  
; Publication No. US20030124537A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Li, Min  
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES  
; FILE REFERENCE: A-70174-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/10/023,208  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/256,163  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Barbarie duck parvovirus  
US-10-023-208-25

Query Match 12.6%; Score 459; DB 14; Length 627;  
Best Local Similarity 26.0%; Pred. No. 6.9e-31;  
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVIFDLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122  
Db 45 WEPTGIWMEHVLNPLWTLADKI-----KNIFQWQFNQDETFDFQLEEGS 93

QY 123 DQGHCHVLIGGKDFGQAQKQWRRQLNVYWSRWLTACNVQTPAERIK---LREIAED 179  
Db 94 EY-IHLHCCI-----AQGNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVLLTYKHQTKDYTKCVLFGNMIAFYFLTKKLISTSPRGGGYFLSDSG 234  
Db 131 QVKIPDFWFSITTKRGQNKVTYAA-----YILHLYLPKKQ-----PELQWAFNTNPLF 180

QY 235 WKTNPLKEGGRHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTTLKELVHK 294  
Db 181 TAAALCLOKQELDLAFQESMAVAVQEDQASTAPL-----ISNRAKQVSNLVDWLIEM 236

QY 295 RVTSPEDDMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354  
Db 237 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIQKDPVLDT 296

QY 355 NFSLPDTRTCRIIFAFHGWNTVYKCHAICCVLNROGGKRNNAVLFHGPASTGKSIIAQAIAQ 414  
Db 297 K-----NRIYQILKLNYPQYGVSLCGWVKEFNKRNAIWLYPATTKTNIAEAIH 351

QY 415 AVGNVGCYNAANVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIRIDOKGKGSK 474  
Db 352 AVFYGCCVNTNENFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIRIDOKGKGSV 411

QY 475 QIBPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530  
Db 412 CIEPTVITNTDMCMVDGNSTTMEHRIPLEERMFQIVLSHLGNFGKISKEVKEP 471

QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568

Db 472 FKWANDNLVPVSEFKVPTNEOTKLTE-----PVPERANEPSEPPKIWAPPTRELEEEI 525  
QY 569 -----VPTPINLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPWPSTPNT 617  
Db 526 LRASPELFASVAPLP-----SSPDTSPKPKTRGEYQVRCAMHSL-DNSMNVFECLEC 577  
QY 618 PVAGTAPTONTGAGSKAC-ODQOLSPTWSEIEEDLRACFGAPLKKDFSE 667  
Db 578 ERANFPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 13  
US-09-792-630-21  
; Sequence 21, Application US/09792630  
; Patent No. US20020168640A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Min  
; APPLICANT: Dahiyat, Bassil I.  
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES  
; FILE REFERENCE: A-70295/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/792,630  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: adeno-associated virus 5  
US-09-792-630-21

Query Match 12.5%; Score 453.5; DB 9; Length 610;  
Best Local Similarity 27.2%; Pred. No. 2e-30;  
Matches 157; Conservative 84; Mismatches 221; Indels 115; Gaps 22;

QY 67 WDQSEDMEWETTVDENTKKQVIFDLSLVKKCLFEVLNTKNIPGDEVNWFVQHEWGKQDQW 126  
Db 35 WEPPSDENLTLVE---QPOLTIVADRIRRVFLTE---WNKFSQESKGFVQFKEGSEY-F 88

QY 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLTACNVQTPAERIKLREI-AEDNEWVT 184  
Db 89 HLHTLVETSGISSMWLGRY-----VSQIRAOQLV---KVVFOGIEPQINDWA 132

QY 185 LLTYKHQTKDYTKCVLFGNMIAFYFLTKKLISTSPRGGGYFLSDSGSWKTNFLKEGE 244  
Db 133 I-----TKVKGKANKVDSGVIPAYLLPK-----VQPELQWAWTNLDE-- 171

QY 245 RHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTTLKE-----LV 292  
Db 172 -YKLAALNLEERKRLVAQFLAESQSQSQAASQ--REFSADPVIKSKTSQKYMALVNLV 228

QY 293 HKRVTSPEDDMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSK 352  
Db 229 EHGITSEKQWLTENKESYLSFNSTGNSRQIKAAALNATKIMSLTKSAVDY----- 279

QY 353 LTNFSLPD---TRTCRIIFAFHGWNTVYKCHAICCVLNROGGKRNNAVLFHGPASTGK 405  
Db 280 LVGSSVPEDISKRNRIWQIFEMNGYDPAYAGSILYGC---QRSFNKRNVTWLYGPATTKG 336

QY 406 SIIAQIAQAGNVGCYNAANVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIR 465  
Db 337 TNIAEAIHTVPPYGCNVNTNENFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIR 396

QY 466 IDQKGGSKQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525  
Db 397 VDQKCHSSVQIDSTPVIIVTSNTNMCVVDGNSTTTFEHQQPLEDEMFKELTKELPDPFGK 456

QY 526 VDKNEWPMTCAMLVKNGYQSTMASYCAKWKGVDPWSE-----NNAEPKVPITPI-- 573  
Db 457 ITKQEVKDFFAW-----AKVQNVPTVTEFKVPVRELACTGKAESKLRPLGD 502

QY 574 -----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600  
Db 503 VTNTYSKLEKRLARLSFVPEPTPRSSDVTVDPAFLRPL 539

```

RESULT 14
US-09-953-351-21
; Sequence 21, Application US/09953351
; Publication No. US20030036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-09-953-351-21

```

Query Match	12.5%	Score 453.5	DB 10	Length 610																
Best Local Similarity	27.2%	Pred. No. 2e+30																		
Matches 157; Conservative	84	Mismatches 221	Indels 115	Gaps 22																
Qy	67	WQOSEDMEWETTVD	MTKQVIFD	SLVKKCLFVLN	TNKFDPGDV	VNMFVQHEWGKDGW	126													
Db	35	WEUPPESDLNLTVE	--OPQ	LTVADRIRV	ELYE---	WNKESKQESKEFV	QPEKGEY-F	88												
Qy	127	HCHVLIGGKDF	-SOAQGK	WRRQLN	YNSRWLV	TACNVQLP	PAERIKLREI	-AEDNEWVT	184											
Db	89	HLHTLVETSGISS	VMVLGRY	-----	V	SQIRALV	--KVV	FQGI	EPQINDWA	132										
Qy	185	LLTYKHQTKDYTK	CVLFGN	MIAYELTK	KIKISTSP	RDGGYFL	SDSGKWTN	FLKEGE	244											
Db	133	I-----	TKV	KGGANKV	DSGYIP	AYLLPK	-----	VQ	PELQAWTNLDE	--	171									
Qy	245	RHLVSKLYTD	DMRPETVETT	TTTAQETK	RGRIO	TKEK	SVIK	ITLKE	-----	LV	292									
Db	172	-YKLAALN	SEERKELVA	QFLAES	QSGEAA	SQ--	REFS	ADPV	IKSTSQ	KYMAV	NVWL	228								
Qy	293	HKRVTSPED	WNMOPDSY	IEEM	AOPG	GENLLK	NTLEICT	LT	LARTK	TAPDL	ILEKAETSK	352								
Db	229	EHGITSEK	QWIOENQ	ESYLSFN	STGNSR	SQIKAA	LDNATK	IMSL	TSK	SAVDY	-----	279								
Qy	353	LTNFSLPD	---T	TCTRIF	AFHGN	--YV	-KV	CHAC	CVLN	RQGR	NAVL	FHGP	PASTGK	405						
Db	280	LVGSSVED	ISKNI	WIQIF	ENMG	YDPAY	AGSIL	GCW	---Q	RSFN	KNTVM	LYGP	ATTGK	336						
Qy	406	SIIAQAI	AAQAVNG	VCYNA	ANVP	FPND	CTN	KNLI	WVEAG	NFGQ	OVNF	KAI	CSG	QOTIR	465					
Db	337	TNIAEA	IAHTVP	FGCV	KNWTN	ENFP	ND	CVDK	MLI	WVEEG	KWTK	WBSA	KIL	GGSKVR	396					
Qy	466	IDQKGK	SQLEPT	PPVIM	TN	TNENIT	TVV	RIG	CEER	PEHT	QPT	RD	RLN	NIHL	THHL	PGDGL	525			
Db	397	VDQCK	SSVQID	STFV	ITSNTN	NCVV	DGNS	TTFE	HQ	QPL	EDRM	FKPEL	TKR	PP	DPFGK	456				
Qy	526	VDKN	EWPMI	CAWL	VKG	YG	QSTMA	SYCA	KWG	KVP	DWSE	-----	NW	AE	PKVPT	PT	--	573		
Db	457	ITKQ	EVKDF	FAW	-----	AK	VQ	VP	VT	HEF	KV	P	REL	AGT	GK	AKSL	KR	PLGD	502	
Qy	574	-----	NL	IGS	ARS	PFT	--	TP	K	STP	LS	QNTYA	-	LTPL	600					
Db	503	VTNTS	YKSL	ERK	ARLS	FV	PE	TP	RS	SD	VT	VD	P	AP	LR	PL	539			

RESULT 15  
US-10-080-376-21  
; Sequence 21, Application US/10080376

```

; Publication No. US20020172968A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080,376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-080-376-21

```

	Query Match	12.5%; Score 453.5; DB 13; Length 610;
	Best local similarity	27.2%; Pred. No. 2e-30;
	Matches 157; Conservative	84; Mismatches 221; Indels 115; Gaps 22;
QY	67	WDQSEDMEWETTVDENMTKKOVFFDSLSVKKCLPELVNTKNIFPGDVNWFVQHEWGKDQGW 126
DB	35	WELPPESDLNLTLVE--QPQLTVADRLRRVFLYE---WNKFESKQESKFFVFQFKGSEY-F 88
QY	127	HCHVLIGGGKDF-SQAQGKWRRRLNVVMSRWLVTACNVQLTPABRIKLREI-AEDNEWWT 184
DB	89	HLHTLVETSGISSMVLGRY-----VSQIRAOQLV--KVVFQGIPEQINDWVA 132
QY	185	LLTYKHQTCKDYTKCVLFGNMLTAYVELTKKKISTSPRDGGYFLLSDSGWKTNFLKERGE 244
DB	133	I-----TKVRKGANKVVDSGYIPAYLLPK-----VQPELOWATNLDB-- 171
QY	245	RHLVSKLYTDMMRPETVETTTVAQETKRGRIOTKKEVSIKNTILKE-----LV 292
DB	172	-YKLAALNLEERKRLVAQFLAESSQRSQEAASQ--REFSADPVIKSTSQKYMALVNWL 228
QY	293	HKRVTSPEDWMNQPDYSIEMMAQPGGENLLKNTLEICTLTATKTAFDLILEKAETSK 352
DB	229	EHGITSEKOWIQBNQESYLSFNSTGNSRSQIKAALEDNATKIMSLTKSAVDY----- 279
QY	353	LTFNFLSPD---TRTCRIFAPFGWN--YV-KVCHAI CVLNRRQGGKENAVLPHGPASTGK 405
DB	280	LVSSVPEDISKRIWIQIFEMWGDYPAYAGSILYGWC---QRSFNKENTVWLYGPATIGK 336
QY	406	SIIAQAIQAQGVNVCYNAANVPFPNDCTNNKLIWVEEAGNFQQQVNFQKAI CSGOTIR 465
DB	337	TNTAEIAHTVPFYGCYNWNTNENPFPPNCDCVDKMLIMWEEGKWTNKVSVESAAILGGSKVR 396
QY	466	IDQKGKSQOIEPTPVIMTNNENITVYRIGCEERPETHOTPRDRMLNIHLTHLPDGDFGL 525
DB	397	VDOCKSSVQIDSTPVTISNTNMCVVVDNGNSTTFEHQQPLEDRMFELTKRPLPPDPGK 456
QY	526	VDKNKPMPMICAWLVKNGYQSTMASYCAKWKGVKPDWSE-----NWAEKPYPTPI-- 573
DB	457	ITQEYKDXDFPW-----AKVQNQVPVTHEFKVPRELAGTKGAEKSLKRPLGD 502
QY	574	-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600
DB	503	VTNTSYKSLEKRLARLSFVPTPRSSDDTVPDPAPLRPL 539

Search completed: January 22, 2005, 04:03:44  
Job time : 92.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:18 ; Search time 22.3333 Seconds  
(without alignments)  
2895.119 Million cell updates/sec

Title: US-10-069-056-14  
Perfect score: 3637  
Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKKDFSEPLNLD 672  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3623	99.6	672	1 UYPVIM	noncapsid protein
2	3556	97.8	721	1 UYPVIM	noncapsid protein
3	3326	91.4	672	1 UYPV1	noncapsid protein
4	3322	91.3	668	1 A44276	noncapsid protein
5	2647	72.8	668	1 UYPVME	noncapsid protein
6	2644	72.7	668	1 UYPVFP	noncapsid protein
7	2642	72.6	668	1 UYPVCP	noncapsid protein
8	2485	68.3	660	1 UYPVPP	noncapsid protein
9	2476	68.1	662	1 UYPVNA	noncapsid protein
10	1621.5	44.6	392	1 UYPVIF	noncapsid protein
11	551	15.1	641	2 S41439	gene NS-1 protein
12	545.5	15.0	641	2 S41434	gene NS-1 protein
13	544	15.0	620	1 UYPVAP	noncapsid protein
14	534.5	14.7	641	2 S41861	gene NS-1 protein
15	459.5	12.6	626	2 S52209	noncapsid protein
16	419.5	11.5	536	1 UYADIA	noncapsid protein
17	350	9.6	726	1 UYPVSI	noncapsid protein
18	343	9.4	671	1 UYPV19	noncapsid protein
19	164	4.5	490	2 T44050	hypothetical prote
20	163.5	4.5	490	2 J01630	noncapsid protein
21	139	3.8	276	2 S26428	hypothetical 31.5K
22	133.5	3.7	849	1 UYPVAD	noncapsid protein
23	132	3.6	545	2 B44054	orf2 protein - Jun
24	124.5	3.4	647	1 W1W139	E1 protein - human
25	120.5	3.3	825	2 G96665	protein F22C12.12
26	116.5	3.2	614	1 W1W141	E1 protein - human
27	114.5	3.1	1098	2 E70232	hypothetical prote
28	113.5	3.1	631	2 S36505	E1 protein - human
29	113.5	3.1	647	2 S36517	E1 protein - human

ALIGNMENTS

RESULT 1  
UYPVIM

noncapsid protein NS1 - minute virus of mice  
C;Species: minute virus of mice, murine parvovirus  
C;Date: 14-Nov-1983 #sequence\_revision 28-Aug-1985 #text\_change 09-Jul-2004  
C;Accession: A03696  
R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.  
Nucleic Acids Res. 11, 999-1018, 1983  
A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.  
A;Reference number: A03696; MUID:83143341; PMID:6298737  
A;Accession: A03696  
A;Molecule type: DNA  
A;Residues: 1-672 <AST>  
A;Cross-references: UNIPROT:P03134; EMBL:V0115  
C;Superfamily: parvovirus noncapsid protein  
C;Keywords: noncapsid protein

Query Match 99.6%; Score 3623; DB 1; Length 672;  
Best Local Similarity 99.7%; Pred. No. 2.2e-255;  
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENNVQLNGKDIGNSYKKSLQDELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENNVQLNGKDIGNSYKKSLQDELKSLQ	60
Qy	61	RGAEITTDQSEDMEWETTVDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW	120
Db	61	RGAEITTDQSEDMEWETTVDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW	120
Qy	121	GKQDQGHCHVLI GSKDFSOAQGWRRRLNYYWSRWLV TACNVQLTPAERIKLREIAEDN	180
Db	121	GKQDQGHCHVLI GSKDFSOAQGWRRRLNYYWSRWLV TACNVQLTPAERIKLREIAEDN	180
Qy	181	EWTLTYKHKQTKDYKCVLFSGNMIAYFLTKKISTSPRDGGYFLSDSGWKTNFL	240
Db	181	EWTLTYKHKQTKDYKCVLFSGNMIAYFLTKKISTSPRDGGYFLSDSGWKTNFL	240
Qy	241	KEGERHLVSKLYTDDMRPETHVTETTTAQTETKGRITQKKEVSIKTTLKELVHKRVTSPE	300
Db	241	KEGERHLVSKLYTDDMRPETHVTETTTAQTETKGRITQKKEVSIKTTLKELVHKRVTSPE	300
Qy	301	DWMMQPDSEIEMMAOPGGENLKNLTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD	360
Db	301	DWMMQPDSEIEMMAOPGGENLKNLTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD	360
Qy	361	TRTCRIFAFHGMNVKVCCHAI CCVLRNQGGRNAVL FHPGASTGKSI IAAIAQAVNGV	420
Db	361	TRTCRIFAFHGMNVKVCCHAI CCVLRNQGGRNVTL FHPGASTGKSI IAAIAQAVNGV	420
Qy	421	CYNAANVPFPNDCTKNLWVEEAGNFGQOVNQKAI CSQGTIRIDOKGKSGKQIEPTP	480
Db	421	CYNAANVPFPNDCTKNLWVEEAGNFGQOVNQKAI CSQGTIRIDOKGKSGKQIEPTP	480

30 112.5 3.1 1008 2 T18832 probable RNA helic  
31 111 3.1 417 2 C70391 hypothetical prote  
32 110.5 3.0 643 2 S36499 E1 protein - human  
33 109 3.0 756 2 T34797 probable ATP bindi  
34 109 3.0 1247 1 VHWVN2 structural polypro  
35 108.5 3.0 497 2 C70454 transcription regu  
36 108.5 3.0 963 2 AD2381 type I site-specif  
37 107 2.9 3678 2 S28916 dyatrophin - mouse  
38 106.5 2.9 405 1 W1WLB4 E1 protein - bovin  
39 106.5 2.9 643 1 S15616 E1 protein - human  
40 106.5 2.9 695 2 T38804 hypothetical WW do  
41 106.5 2.9 3848 2 T17414 TipC protein - eli  
42 106 2.9 474 2 T00943 hypothetical prote  
43 106 2.9 526 2 F85086 hypothetical prote  
44 106 2.9 603 1 W1WL8 E1 protein - human  
45 106 2.9 691 1 TVVPAS large T antigen -

QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540  
D 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540  
QY 541 NGYQSTWASACAKGKVPDWSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600  
D 541 NGYQSTWASACAKGKVPDWSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600  
QY 601 ASDLEDLALPWSPTNTPVAGTAETONTGEAGSKACODGQSLPTWSEIEDLRACFGAEP 660  
D 601 ASDLEDLALPWSPTNTPVAGTAETONTGEAGSKACODGQSLPTWSEIEDLRACFGAEP 660  
QY 661 LKQDFSEPLNLD 672  
D 661 LKQDFSEPLNLD 672

RESULT 2  
UYPVIM  
noncapsid protein NS1 - minute virus of mice (strain MVMi)  
C:Species: minute virus of mice, murine parvovirus  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
C:Accession: A23008; A29510  
R:Sanli, R.; McMaister, G.K.; Hirt, B.  
Nucleic Acids Res. 13, 3617-3633, 1985  
A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous  
A:Reference number: A23008; UID:85242059; PMID:3855242  
A:Accession: A23008  
A:Molecule type: DNA  
A:Residues: 1-721 <SAH>  
A:Cross-references: UNIPROT:Q84363; EMBL:X02481  
R:Astelli, C.R.; Gardiner, E.M.; Tattersall, P.  
J. Virol. 57, 656-669, 1986  
A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and  
A:Reference number: A29510; UID:86115415; PMID:3502703  
A:Accession: A29510  
A:Molecule type: DNA  
A:Residues: 1-645, '1', 647-721 <AST>  
A:Cross-references: EMBL:M12032; NID:g332289; PIDN:AAA69566.1; PID:g825477  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 97.8%; Score 3556; DB 1; Length 721;  
Best Local Similarity 97.6%; Pred. No. 1.8e-250;  
Matches 656; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKEKSNQEVFSFVKNNVQLNGKDIGWNSYKKELODELKSLQ 60  
D 50 MAGNAYSDEVILGATNWLKEKSNQEVFSFVKNNVQLNGKDIGWNSYKKELODELKSLQ 109  
QY 61 RGAETTTDQSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120  
D 110 RGAETTTDQSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIFFPGDVNWFVQHEW 169  
QY 121 GKDGQWHCHVLIGKDFSQAGKQWRRQNLNYSRWLVLTACNVQLTPAERIKLREIAEDN 180  
D 170 GKDGQWHCHVLIGKDFSQAGKQWRRQNLNYSRWLVLTACNVQLTPAERIKLREIAEDS 229  
QY 181 EWTLLTYKHQTKKDYKCVLFQNMIAFYFLTKKKISTSPPRDGGYFLSSDSGKTNFL 240  
D 230 EWTLLTYKHQTKKDYKCVLFQNMIAFYFLTKKKISTSPPRDGGYFLSSDSGKTNFL 289  
QY 241 KEGERHLVSKLYTDDMEPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300  
D 290 KEGERHLVSKLYTDDMEPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 349  
QY 301 DMMMQPDSYIEMMAQPGGNNLKNLTLEICTLTARTKTAFLDILKEAETSCLTNFSLPD 360  
D 350 DMMMQPDSYIEMMAQPGGNNLKNLTLEICTLTARTKTAFLDILKEAETSCLTNFSLPD 409  
QY 361 TRTCRIFAEGHNVYKVCCHAICCVLNROGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG 420

D 410 TRTCRIFAEGHNVYKVCCHAICCVLNROGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG 469  
QY 421 CYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAICSGQTIIRDQKSGSKQIEPTP 480  
D 470 CYNAANVPFNDCTNKNLIWVEEAGNFGQVNOFKAICSGQTIIRDQKSGSKQIEPTP 529  
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540  
D 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 589  
QY 541 NGYQSTWASACAKGKVPDWSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600  
D 590 NGYQSTWASACAKGKVPDWSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 649  
QY 601 ASDLEDLALPWSPTNTPVAGTAETONTGEAGSKACODGQSLPTWSEIEDLRACFGAEP 660  
D 650 ASDLEDLALPWSPTNTPVAGTAETONTGEAGSKACODGQSLPTWSEIEDLRACFGAEP 709  
QY 661 LKQDFSEPLNLD 672  
D 710 LKQDFSEPLNLD 721

RESULT 3  
UYPVV1  
noncapsid protein NS1 - parvovirus H1  
C:Species: parvovirus H1  
A:Note: host Homo sapiens (man)  
C:Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 09-Jul-2004  
C:Accession: A03695  
R:Rhode III, S.L.; Paradiso, P.R.  
J. Virol. 45, 173-184, 1983  
A:Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid  
A:Reference number: A03695; UID:83112183; PMID:6823009  
A:Accession: A03695  
A:Molecule type: DNA  
A:Residues: 1-672 <RHO>  
A:Cross-references: UNIPROT:P03133; EMBL:X01457; NID:g60993; PIDN:CAA25689.1; PID:g60994.1  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 91.4%; Score 3326; DB 1; Length 672;  
Best Local Similarity 91.1%; Pred. No. 8.7e-234;  
Matches 612; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKEKSNQEVFSFVKNNVQLNGKDIGWNSYKKELODELKSLQ 60  
D 1 MAGNAYSDEVILGATNWLKEKSNQEVFSFVKNNVQLNGKDIGWNSYKKELODELKSLQ 60  
QY 61 RGAETTTDQSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120  
D 61 RGAETTTDQSEDMEWESA VDDMTKKQVFIYDLSLVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120  
QY 121 GKDGQWHCHVLIGKDFSQAGKQWRRQNLNYSRWLVLTACNVQLTPAERIKLREIAEDN 180  
D 121 GKDGQWHCHVLIGKDFSQAGKQWRRQNLNYSRWLVLTACNVQLTPAERIKLREIAEDS 180  
QY 181 EWTLLTYKHQTKKDYKCVLFQNMIAFYFLTKKKISTSPPRDGGYFLSSDSGKTNFL 240  
D 181 EWTLLTYKHQTKKDYKCVLFQNMIAFYFLTKKKISTSPPRDGGYFLSSDSGKTNFL 240  
QY 241 KEGERHLVSKLYTDDMEPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300  
D 241 KEGERHLVSKLYTDEMKEPETVETTTTAQEAARGRIQTRVEVSIKTTLKELVHKRVTSPE 300  
QY 301 DMMMQPDSYIEMMAQPGGNNLKNLTLEICTLTARTKTAFLDILKEAETSCLTNFSLPD 360  
D 301 DMMMQPDSYIEMMAQPGGNNLKNLTLEICTLTARTKTAFLDILKEAETSCLTNFSLPD 360  
QY 361 TRTCRIFAEGHNVYKVCCHAICCVLNROGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG 420  
D 361 TRTCRIFAEGHNVYKVCCHAICCVLNROGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG 420



[illegible]

421 CYNAANVPPFPNDCTNNKNIWIWEAGNFGQVQNOFKKACSGQTTIRIDQKGKSGKQIEPTP 480

481 VIMITNENITVVRIGCEERPEHTQPIRDRLMLNIHLTHHLPDGLVDKNEWPMICAWLVK 540

481 VIMITNENITVVKIGCEERPEHTQPIRDRLMLNIHLTHHLPDGLVDKNEWPMICAWLVK 540

541 NGYQSTMASYCAKWKGVDPWSENWABPKVPTPINLLGSARSPTFTPKSTPISQNYALTP 600

541 NGYQSTMASYCAKWKGVDPWSENWABPKVPTTEINSVGSTNSP--SPKSTPLSQNYALTP- 597

601 ASDLEDLALPMSHTNTPTVAGTAETONTCEAGSKACODGOLSPTWSETEEDLRACFGAEP 660

598 -SDLEDLALPMSHTNTPTVAGTAETONTCEAGSKACODGOLSPTWSETEEDLRACFSSEH 656

661 LKQDFSEPLNLD 672

657 WKSDEQLPNLD 668

RESULT 5

UYPVME

noncapsid protein NS1 - mink enteritis virus (strain Abashiri)

C/Species: mink enteritis virus, MEV

C/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004

C/Accession: A38350

R/Kariatsunari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shing

J. Gen. Virol. 72, 867-875, 1991

A/Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the

A/Reference number: A38350; MUID:91202123; PMID:2016597

A/Accession: A38350

A/Molecule type: DNA

A/Residues: 1-668 <KAR>

A/Cross-references: UNIPROT:P27438; GB:D00765; NID:G222435; PIDN:BA00662.1; PID:G222436

C/Superfamily: parvovirus noncapsid protein

C/Keywords: noncapsid protein

Query Match 72.8%; Score 2647; DB 1; Length 668;

Best Local Similarity 73.1%; Pred. No. 2.1e-184;

Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVLGATNWLKESKQEVPSFVKPNENVOLNGKDIGWNSYKKELOEBELKSIQ 60

Db 1 MSGNQYTEBEVMEGVNWLKHAENAEFSFVKCDNVQLNGKDVHNNYTKPQNEBELSLI 60

Qy 61 RGAETTWQDS--EDMEWETTVDGMTKKQVIFDSLKVKCLFVNLTKNIFPGDVNMFVQH 118

Db 61 RGAQTAMDQTEEEEDMDESEVDSLAKQVQTFDALLKCLFEVFSKNIEPNECVWFQIH 120

Qy 119 EWGKDQGHCHVLIIGKDFSQAGKWRRLQNLVYWRSLVTACNVQLTAPARIKLREIAE 178

Db 121 EWGKDQGHCHVLLSHSKNLQATGKWLRRQMNYWSRWLVTLCSVNLTPTEKIKLREIAE 180

Qy 179 DNEWTLITYKHQKQKDYTKCVLFQNMAYYFLTKKIKISTSPRDGGYFPLSSDSGKTN 238

Db 181 DSEWTLITYRHQKQKDYVVMVHFGNMAYYFLTKKIKI-VHMTKESGYFLSTDSGWKFN 239

Qy 239 FLKEGERHLSKLYDDMRPEVTVTYTAQETKRGRIQTKKEVSIKTLKELVHKRVTS 298

Db 240 FMKQDRHTVSTLYTEQMKPEVTVTYTAQETKRGRIQTKKEVSIKCTLDLVSQKRVTS 299

Qy 299 PEDMMWQPDSEYEMWQPGENLLKNTLEICTLTARTKTAFOILLEAKTSKLTNFSL 358

Db 300 PEDMMWQPDSEYEMWQPGENLLKNTLEICTLTARTKTAFOILLEAKNTKLTNFDL 359

Qy 359 PDTRTCRIFAPHGNVYKVCIAICCVLRQGGKRNVLPHGPASTGKSIIAQATAQAVN 418

Db 360 ANSTRCQIFRMHGWNIKVCHAIACVNLNQGKRNVLPHGPASTGKSIIAQATAQAVN 419

Qy 419 VGCYNAAVNPFNDCTNNKNIWIWEAGNFGQVQNOFKKACSGQTTIRIDQKGKSGKQIEP 478

Db 420 VGCYNAAVNPFNDCTNNKNIWIWEAGNFGQVQNOFKKACSGQTTIRIDQKGKSGKQIEP 479

Qy 479 TPIVMTNTNITVVRIGCEERPEHTQPIRDRLMLNIHLTHHLPDGLVDKNEWPMICAWL 538

Db 480 TPVIMTNNENITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDRKEEPLICAWL 539  
Qy 539 VKNGYQSTWASYCAKWKGVDPWSNWAEPKVPPTPINLLGSARSPTTTPKSTPLSONVALT 598  
Db 540 VKHGYSQSTWANYTHHWKGVPEWDENWAEPKIQEGL-ISPCKDLETAQASNPQSDHVL 598  
Qy 599 PLASDLLEALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPPTWSEIEEDLRACFGA 658  
Db 599 PLTDPDVVDLALPWSPTDTPIAETA-NQOSNQLG-VTHKDVQASPTWSEIEADLRIFT 656  
Qy 659 EPLKDFSEPLN 670  
Db 657 EQLEEDFRDLD 668

## RESULT 6

UYPVP noncapsid protein NS1 - feline panleukopenia virus (strain 193)  
C:Species: feline panleukopenia virus, FPLV  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: A36608  
R:Martyn, J.C.; Davidson, B.E.; Studdert, M.J.  
J. Gen. Virol. 71, 2747-2753, 1990  
A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus  
A:Reference number: A36608; MUID:91073139; PMID:2174965  
A:Accession: A36608  
A:Molecule type: DNA  
A:Residues: 1-668 <MAR>  
A:Cross-references: UNIPROT:P24842; GB:X55115; NID:G60863; PIDN:CAA38910.1; PID:G60864  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 72.7%; Score 2644; DB 1; Length 668;  
Best Local Similarity 73.1%; Pred. No. 3.4e-184;  
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;  
Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60  
Db 1 MSGNQYTEEVGVNWLKHAENAEAFVFKCDNVQLNGKDVRRNNYTKPIQNEELTSLI 60  
Qy 61 RGAETWDS--EDMEWETTVDKTKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNFWQH 118  
Db 61 RGAQTAMDQTEEBEEMDESEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECVWFIQH 120  
Qy 119 EWGKDGQWCHVLIGKDFSOAGKWRRLNVYWSRWLVACNVLTTPAERIKLREIAE 178  
Db 121 EWGKDGQWCHVLHSLKQLQATGKWLRRQNMWYWSRWLVTLCSVNLTPTEKIKLREIAE 180  
Qy 179 DNEWVTLTYKHQTKKDYTKCVLFGNMIAIYFLTKKKISTSPRDGGYFLSSDSGKTN 238  
Db 181 DSEWVTLTYRHKQTKDYVKVWFHFGNMIAIYFLTKKKI-VHMTKESGYFLSDSGWKN 239  
Qy 239 FLKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTS 298  
Db 240 FMKYQDRHTVSTLYTEQMKPEVETVTTTAQETKRGRIQTKKEVSIKTLRDLVSKRVTS 299  
Qy 299 PEDMMMLQPSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 358  
Db 300 PEDMMMLQPSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 359  
Qy 359 PDTRTCRIFAFHGMNYYKVCCHAICCVLNROGGRNAVLPHGPASTGKSIQAIAQAVGN 418  
Db 360 ANSRCTQIFRMHGMNWKVCHAIACVNLROGGRNTVLFHGPASTGKSIQAIAQAVGN 419  
Qy 419 VGCYNAANVPFNDCTNKNLIWVEEAGNFQQVNOQFKAICSGQTIIRIDQKGGSKQIEP 478  
Db 420 VGCYNAANVPFNDCTNKNLIWVEEAGNFQQVNOQFKAICSGQTIIRIDQKGGSKQIEP 479  
Qy 479 TPVIMTNNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDRKEEPLICAWL 538  
Db 480 TPVIMTNNENITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDRKEEPLICAWL 539

Qy 539 VKNGYQSTWASYCAKWKGVDPWSNWAEPKVPPTPINLLGSARSPTTTPKSTPLSONVALT 598  
Db 540 VKHGYSQSTWANYTHHWKGVPEWDENWAEPKIQEGL-ISPCKDLETAQASNPQSDHVL 598  
Qy 599 PLASDLLEALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPPTWSEIEEDLRACFGA 658  
Db 599 PLTDPDVVDLALPWSPTDTPIAETA-NQOSNQLG-VTHKDVQASPTWSEIEADLRIFT 656  
Qy 659 EPLKDFSEPLN 670  
Db 657 EQLEEDFRDLD 668

## RESULT 7

UYPVCP noncapsid protein NS1 - canine parvovirus (strain N)  
C:Species: canine parvovirus, CPV  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C:Accession: A29962  
R:Reed, A.P.; Jones, E.V.; Miller, T.J.  
J. Virol. 62, 266-276, 1988  
A:Title: Nucleotide sequence and genome organization of canine parvovirus.  
A:Reference number: A29962; MUID:88062992; PMID:2824850  
A:Accession: A29962  
A:Molecule type: DNA  
A:Residues: 1-668 <REE>  
A:Cross-references: UNIPROT:P12929; EMBL:M19296; NID:G333436; PIDN:AAA67459.1; PID:G33343  
C:Superfamily: parvovirus noncapsid protein  
C:Keywords: noncapsid protein

Query Match 72.6%; Score 2642; DB 1; Length 668;  
Best Local Similarity 73.1%; Pred. No. 4.8e-184;  
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;  
Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60  
Db 1 MSGNQYTEEVGVNWLKHAENAEAFVFKCDNVQLNGKDVRRNNYTKPIQNEELTSLI 60  
Qy 61 RGAETWDS--EDMEWETTVDKTKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNFWQH 118  
Db 61 RGAQTAMDQTEEBEEMDESEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECVWFIQH 120  
Qy 119 EWGKDGQWCHVLIGKDFSOAGKWRRLNVYWSRWLVACNVLTTPAERIKLREIAE 178  
Db 121 EWGKDGQWCHVLHSLKQLQATGKWLRRQNMWYWSRWLVTLCSVNLTPTEKIKLREIAE 180  
Qy 179 DNEWVTLTYKHQTKKDYTKCVLFGNMIAIYFLTKKKISTSPRDGGYFLSSDSGKTN 238  
Db 181 DSEWVTLTYRHKQTKDYVKVWFHFGNMIAIYFLTKKKI-VHMTKESGYFLSDSGWKN 239  
Qy 239 FLKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTS 298  
Db 240 FMKYQDRQIVSTLYTEQMKPEVETVTTTAQETKRGRIQTKKEVSIKTLRDLVSKRVTS 299  
Qy 299 PEDMMMLQPSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 358  
Db 300 PEDMMMLQPSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 359  
Qy 359 PDTRTCRIFAFHGMNYYKVCCHAICCVLNROGGRNAVLPHGPASTGKSIQAIAQAVGN 418  
Db 360 ANSRCTQIFRMHGMNWKVCHAIACVNLROGGRNTVLFHGPASTGKSIQAIAQAVGN 419  
Qy 419 VGCYNAANVPFNDCTNKNLIWVEEAGNFQQVNOQFKAICSGQTIIRIDQKGGSKQIEP 478  
Db 420 VGCYNAANVPFNDCTNKNLIWVEEAGNFQQVNOQFKAICSGQTIIRIDQKGGSKQIEP 479  
Qy 479 TPVIMTNNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDRKEEPLICAWL 538  
Db 480 TPVIMTNNENITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPDGLVDRKEEPLICAWL 539  
Qy 539 VKNGYQSTWASYCAKWKGVDPWSNWAEPKVPPTPINLLGSARSPTTTPKSTPLSONVALT 598  
Db 540 VKHGYSQSTWANYTHHWKGVPEWDENWAEPKIQEGL-ISPCKDLETAQASNPQSDHVL 598

[illegible]

Db 421 YNAANVFPNDCTNKLWIEEAGNSQVNOQKAIKCSGTIRIDQKSGSKQIEPTPV 480  
Qy 482 IMTNENITVIRICERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVKN 541  
Db 481 IMTNEDITKIRICERPEHTQPIRDRMLNINUTRKLPGDGLLETEWPLICAWLVKK 540  
Qy 542 GYSTMASYCAKMGKVPDWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSONYALTPLA 601  
Db 541 GYQATMASYHHWGNVDPWSEKWEKVPKQPTINTPTDSQIS-TSVKTSPADNNVYAATPIQ 599  
Qy 602 SDLE-DLALPSPWTPNTPVAGTAETQNTGAGSKACODGQI-----SPTWSEIEDLRACFG 657  
Db 600 EDLDLALPSPWSEPTTPTTNLHLTTP-----PDSAIRTPSPWSEIETDIRACFG 652

RESULT 10  
UYPVIF  
noncapsid protein NS1 - feline panleukopenia virus (fragment)  
C/Species: feline panleukopenia virus, FPLV  
C/Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C/Accession: A03697  
R/Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.  
J. Virol. 55, 574-587, 1985  
A/Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus  
A/Reference number: A03697; MUID:85265017; PMID:2991581  
A/Accession: A03697  
A/Molecule type: DNA  
A/Residues: 1-392 <CAR>  
A/Cross-references: UNIPROT:P06431; EMBL:M10824; NID:g333474; PIDN:AAA47160.1; PID:g333474  
C/Superfamily: parvovirus noncapsid protein  
C/Keywords: noncapsid protein

Query Match 44.6%; Score 1621.5; DB 1; Length 392;  
Best Local Similarity 77.5%; Pred. No. 3.6e-110; Indels 3; Gaps 3;  
Matches 306; Conservative 33; Mismatches 53;  
Qy 276 IQTKKESVIKTLKELVHKRVTSPEWMMQPSYIEMMAQPGGENLLKNTLICTLTLA 335  
Db 1 IQTKKESVIKTLRLDLSKRVTSPEWMMQPSYIEMMAQPGGENLLKNTLICTLTLA 60  
Qy 336 RTKTAFLILEKARTSKLTNFSLPDTRTCIFAFHGNVYKVAICCVLNROGKKNV 395  
Db 61 RTKTAFLILEKANNTKLTNFDLANSRTQIFRMHGNWIKVCHAIACVLNROGKKNV 120  
Qy 396 LFHGPASTGKSIQAIAQAVGNVCYNAANVPFNDCTNKLWIEEAGNFGQVNOF 455  
Db 121 LFHGPASTGKSIQAIAQAVGNVCYNAANVPFNDCTNKLWIEEAGNFGQVNOF 180  
Qy 456 KAICSGQTIRIDQKSGSKQIEPTPVIIMTNENITVIRIGCEERPEHTQPIRDRMLNIHL 515  
Db 181 KAICSGQTIRIDQKSGSKQIEPTPVIIMTNENITVIRIGCEERPEHTQPIRDRMLNIHL 240  
Qy 516 THLPGDFGLVDKNEWPMICAWLVKNGYQSTMASYCAKMGKVPDWSENWAEKVPPTPINL 575  
Db 241 VCKLPGDFGLVDKEWPLICAWLVKHYESTMANVTHHWGKVPEDENWAEKPIQEGVNS 300  
Qy 576 LGSARSPFTTPKSTPLSONYALTPLASDLELALPSPWTPNTPVAGTAETQNTGAGSKA 635  
Db 301 PG-CKOLETQASNPQSDHVLTPITPDVVDLALPSPWTPNTPVAGTAETQNTGAGSKA 357  
Qy 636 CQDQLSPTWSEIEDLRACFGAPLKKDFSEPLN 670  
Db 358 HKDVQASPTWSEIEDLRACFGAPLKKDFSEPLN 670

RESULT 11  
S41439  
gene NS-1 protein - Aleutian mink disease virus  
C/Species: Aleutian mink disease virus  
C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C/Accession: S41439  
R/Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.

submitted to the EMBL Data Library, January 1994  
A/Description: Sequence comparison of the non-structural genes of four different types of  
A/Reference number: S41434  
A/Accession: S41439  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-641 <GOT>  
A/Cross-references: UNIPROT:Q65017; EMBL:229576  
C/Genetics:  
A/Introns: 586/1  
C/Superfamily: parvovirus noncapsid protein

Query Match 15.1%; Score 551; DB 2; Length 641;  
Best Local Similarity 25.8%; Pred. No. 5.4e-32;  
Matches 178; Conservative 108; Mismatches 262; Indels 142; Gaps 25;

Qy 18 KEKSQEVFSFVKENNVQLNGKDIGWNSYKKELODE-----LKSLQGAETWTD 68  
Db 8 EQRRLOQLYVOLKKEIN--DGEGVAVLFOQKYTYTDKDKPTKATPLRTTSSDLRLAFD 64  
Qy 69 QSDMEWETTVDKTK-----KOVFIPLDSLVKKCLFEVLNTKNIFFPGDVNMFVQ 117  
Db 65 SISE-NITASNEQLTNGINFCKLTGLKTLTLLLDKHVSHRWD--NNK-----VNLIIWQ 115  
Qy 118 HEWKGQGHCHVILG-----GKDFSOQAGKWRRLN-----VWWSRLVTACNVQ- 164  
Db 116 IEKGTQOQHHCCLGYFDRNEDPKVQKSLG-WFMKRLNKLIALIYSNH-----HCDIOD 170  
Qy 165 -LTPAERIKLRETAEDNEWVTLTYKHKOTKDYTKVLFGNMIAVYFLTKKISTSPPR 223  
Db 171 IKDPEKAKMLKVVSDGGTPKPKYFNKQTKQYKPVHLRDFYFIYLFNKKINDTSM- 229  
Qy 224 DGGYFLSSDSGKWTNFKEGEHLVSKLYTDDMRPETVTV-----TTAQ 269  
Db 230 -DCYFAAGNGIVDN-LTNKERKTLKMYLDEQSSDIMDANIDWEDQDAPKVTDOITSA 287  
Qy 270 ETKRG-----RIQTKKEV-----SIKTLKELVHKRVTSPEWMMQ 307  
Db 288 TTKTGTSLIWKSCATKVTSKKEVANFPVQPSKKLYSAQNTLDALFNVGCTPEDMIKQS 347  
Qy 308 DSVIEMMAQPGGENLLKNTLICTLARTKTAFDLILE--KAETSKLTNFSLPDTRTCR 365  
Db 348 KYLELSLENGPKQKINTLLHNNQVKTSTMTAFDCCIKNEDDDKPLLATIKDM---- 403  
Qy 366 IFAFHGNVYKVAICCVLNROGKKNVLFHGPASTGKSIQAIAQAVGNVCYNA 425  
Db 404 -----GLNEQYLKVLCTILTQGGKKGCIWFYGGTGKTLASLICKATVNGVMTS 458  
Qy 426 NVNFPNDCTNKLWIEEAGNFGQVNOFKAICSGQTIRIDQKSGSKQIEPTPIMTT 485  
Db 459 NPNFPWTDGNNRIIWAEECNLGNVVEDFKAITGGDVKVDTKNQPSIKGC-VIVTS 517  
Qy 486 NENITVIRIGCEERPEHTQPIRDRMLNIHL-----THHLPDGLVDKNEWPMICAWL 538  
Db 518 NTNITKVTGCVETNAHAEPKMKIRCMKNTINPKTITPG-----MLKRWL 566  
Qy 539 VKNGYQSTMASYCAKMGKVP-DWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSONYAL 597  
Db 567 -----NTWDRQPIQLSHEMPELYLET---TGPNSSATTATKNTGNSQ---- 605  
Qy 598 TPLASDLELALPSPWTPNTPVAGTAETQN 627  
Db 606 PTTAKAESVNTENCDDTPKRGASVPPKQH 635

RESULT 12  
S41434  
gene NS-1 protein - Aleutian mink disease virus  
C/Species: Aleutian mink disease virus  
C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C/Accession: S41434  
R/Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.  
submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of  
A;Reference number: S41434  
A;Accession: S41434  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-641 <GOT>  
A;Cross-references: UNIPROT:Q65020; EMBL:Z29577  
C;Genetics:  
A;Insertions: 586/1  
C;Superfamily: parvovirus noncapsid protein

Query Match 15.0%; Score 545.5; DB 2; Length 641;  
Best Local Similarity 25.2%; Pred. No. 1.3e-31;  
Matches 182; Conservative 110; Mismatches 260; Indels 171; Gaps 26;

Qy 18 KEKSNQEVFVFKNENVLQNGDKIGWNSYKKEQLQEDS-----LKSLQSGAETWD 68  
Db 8 EQRLQLSFE-KPKTE--VADGGLAWLFOQKTYTDKDKPKTKATPLPRTTSSDLRLAFD 64  
Qy 69 QSEDMETVDEMTKKQVFIQSLVKKCLFEVLNTKNI-----FPGDVNMFVQHEWKGKD 123  
Db 65 SIEE-TLKTSNQCLTNNDINFCFLTKLGL--VLLDKHVRSHRDANKVFIQVEKGKT 121  
Qy 124 QGWHCHVLIG-----GKDFSAQAGKWRRLN-----VYMSRWLVTAACVOLTPAERIK 172  
Db 122 QQFHHCHCLGYFDKDESDKDVQKSLG-WFIKGLNKLDAVINSNHHCDIOGQSDGRADN 180  
Qy 173 LREIAEDNEWTLTYKHKOTKKDYTKCVLFGNMIAYYFLTKKISTSPRDG--GYFLS 230  
Db 181 LKTIWEDGP-TKPKYFNKQTKQDYNKPTNLRDYTLIYLFNKDKIT-----QEGMDGYAS 235  
Qy 231 SDGKWTNFKLKEGRHVLVSKLYTDDM-----RPETVETTTVAQETPKRG 274  
Db 236 GNGGIIDN-LTNKRALRKWYLDQSQGILDEDIDWEDSQSAPKVTDDQTSATSKTGS 294  
Qy 275 -----RQTKKEVSI-----KTLTKELVHKRVTSPEMMQDPSYIEMM 314  
Db 295 LVMSKCATKVTSKVEAIPVKQPSKQWTSQNTLDDLYMFCFTPEDMILKMSDRYLEMS 354  
Qy 315 AQPGENLLKNTLEICTLTARTKTAFLDILE--KAETSKLTNFSLPDTRTCRIFAPHGW 372  
Db 355 LEFNGAQKINTLLHMNOVRTSSNLTAPECILKNEEDEDKPLDTIKDM-----GL 405  
Qy 373 NYKVKCHAICCVLNRQGGKRNALFHGPASTGKSIIAQIAQAVGNVGCYNAANVPFN 432  
Db 406 NEQHLKVLCTILTQSGKRGCVWFYFGGTGKTLASLICKAVNYGMVWTSNPNPMT 465  
Qy 433 DCTNKNLIWVEEAGNFQGVNQVQKAIKCSGQTIKIDQKGSQKIEPTFVIMTNENITV 492  
Db 466 DCGNRNIWAEECGNIGNYVEDEKAITGGGVKVDTKNKPQSIKGS-VIVTSNTNITKV 524  
Qy 493 RIGCEERPEPTQIRDRMLNIHL-----THHLPDGLVDKNWPMICAWLVGNGYQS 545  
Db 525 TVGCVETNVIAEPLKQRMVKIRCMKPVNPKVTGP-----MLKTI----- 566  
Qy 546 TMASYCAKWKGVDPWSENWAEKVPVTPINLLGSARSPTTPKSTPLSQNYALTPASDLE 605  
Db 567 -----STWDRIIP-----IKLSHEMP 581  
Qy 606 DLALPEWSTNTVAGTAETQNTGEGSKACQDQGLSPTWSEIEEDLRACFGAEPKQF 665  
Db 582 ELYLET-SGPNIS-SATTATKST-----GNLQPTTAETAESVNTANCDDTPKRGAS 628  
Qy 666 SEP 668  
Db 629 SVP 631

RESULT 13  
UYVPAP  
noncapsid protein NS1 - Aleutian mink disease virus (strain ADV-G)  
N;Alternate names: left-ORF protein  
C;Species: Aleutian mink disease virus

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: A36760; A35529  
R;Bloom, M.E.; Alexandersen, S.; Perryman, S.; Lechner, D.; Wolfinbarger, J.B.  
J. Virol. 62, 2903-2915, 1988  
A;Title: Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus  
A;Reference number: A36760; MUID:88275062; PMID:2839709  
A;Accession: A36760  
A;Molecule type: DNA  
A;Residues: 1-620 <BL2>  
A;Cross-references: UNIPROT:P24030; EMBL:M20036  
C;Superfamily: parvovirus noncapsid protein  
C;Keywords: noncapsid protein

Query Match 15.0%; Score 544; DB 1; Length 620;  
Best Local Similarity 26.6%; Pred. No. 1.7e-31;  
Matches 157; Conservative 104; Mismatches 229; Indels 100; Gaps 18;

Qy 18 KEKSNQEVFVFKNENVLQNGDKIGWNSYKKEQLQEDS-----FQDVNMFVQHEWKGQGW 126  
Db 38 EQRLQLVYVOLKKEIN---DGGVAVLFOQKTYTDKDKPKTKATPLPRTTSSDLRLAFD 94  
Qy 77 TTVDGMTKKQVFIQSLVKKCLFEVLNTKNI-----FQDVNMFVQHEWKGQGW 126  
Db 95 SIEENLTASNEHLTNNEINFCKLTGLKTLILLIDKHVKSHEWDSKNVLIWQIEKGTQOF 154  
Qy 127 HCHVLIG-----GKDFSAQAGKWRRLN-----VYMSRWLVTAACVQ--LTPAERIKL 173  
Db 155 HIHCLGLGYFDKDESDKDVQKSLG-WFMKRLNKLDAVINSN---HHCDIQIDKPEDRAKN 210  
Qy 174 RETAEDNEWTLTYKHKOTKKDYTKCVLFGNMIAYYFLTKKISTSPRDG--GYFLS 233  
Db 211 LKWIWEDGPTKPKYFNKQTKQDYNKPVHLRDYTFIYLFNKDKINTDSM--DGYFAAGNG 268  
Qy 234 GWTNFKLKEGRHVLVSKLYTDDMRPETVTV-----TTAQTETRG----- 274  
Db 269 GIVDN-LTNKERTKRLKMYLDEQSSDIMDANIDWEDQDAPKYVDQDTSATTKGTSLIW 327  
Qy 275 -----RQTKKEV-----SIKTTTLKELVHKRVTSPEMMQDPSYIEMMAQP 317  
Db 328 KSCATKVTSKVEANPVQPSKLYSAQSTLDALFNVGCTPEDMIKQSDKYLESLPEP 387  
Qy 318 GGENLLKNTLEICTLTARTKTAFLDILE--KAETSKLTNFSLPDTRTCRIFAPHGWYV 375  
Db 388 NGPQKINTLLHMNOVRTSSNLTAPECILKNEEDEDKPLLATIKDM-----GLNEQ 438  
Qy 376 KVCHAICCVLNRQGGKRNALFHGPASTGKSIIAQIAQAVGNVGCYNAANVPFNPDCT 435  
Db 439 YLKKVLCITILTQSGKRGCVWFYFGGTGKTLASLICKAVNYGMVWTSNPNPMTDCG 498  
Qy 436 NKXLIWVEEAGNFQGVNQVQKAIKCSGQTIKIDQKGSQKIEPTFVIMTNENITVVRIG 495  
Db 499 NRNIWAEECGNIGNYVEDEKAITGGGVKVDTKNKPQSIKGC-VIVTSNTNITKVTVG 557  
Qy 496 CEEPRPEPTQIRDRMLNIHL-----THHLPDGLVDKNWPMICAWL 538  
Db 558 CVETNAHAEPKQRMKIRCMKTPINPKTITGP-----MLKRWL 596

RESULT 14  
S41861  
gene NS-1 protein - Aleutian mink disease virus  
C;Species: Aleutian mink disease virus  
C;Date: 15-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C;Accession: S41861; S41436  
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.  
submitted to the EMBL Data Library, January 1994  
A;Description: Sequence comparison of the non-structural genes of four different types of  
A;Reference number: S41434  
A;Accession: S41861  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-641 <GOT>  
A;Cross-references: UNIPROT:Q65023; EMBL:X77085; NID:9452597; PIDN:CAA54355.1; PID:945252

## C;Genetics:

A;Introns: 586/1

C;Superfamily: parvovirus noncapsid protein

Query Match 14.7%; Score 534.5; DB 2; Length 641;  
Best Local Similarity 25.7%; Pred. No. 8.5e-31;  
Matches 178; Conservative 106; Mismatches 250; Indels 159; Gaps 27;

QY 18 KEKSNQEVFSVFVFNENVLQNGKIDGWSYKKELEQDE-----LKSQRGAETWTD 68  
DB 8 EOKRLQDLYTLQKKE---VADGEGLAWLFQKTYTDKDKNPKTKATPRLTSSDLRLAFD 64  
QY 69 QSEDM-----EWETTVEMT-----KKQVFIFDSLVKCKLFEVLNTKNIPFGDVNWFVQ 117  
DB 65 SIEETLIASNEW-LTKDEINFCKLTGLGKTLVLVDKVKSHRWA-----DKINFIQ 115  
QY 118 HEWKQDQWHCHVLIG-----GKFSQAQGWRRQLNVYNSRWLVTACNVOLTPAER 170  
DB 116 IEKGTQHFHHCCLGYFDKNEPKDVQKSLG-WLIKKLN-----RDAAIFSNHHCQDQD 170  
QY 171 IKLREIAEDN-EWV-----TLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPPR 223  
DB 171 IKDPEAKANNLKWIEDGTPKPYKYNNKQTKQDYNKPVTLSDYVILYFNKDKIH-----K 226  
QY 224 DG--GYFLSSDSGKWTNFKEGERHLVSKLYTDDMRPETVE-----TTVTT 267  
DB 227 EGMGYAAGNGGLIDN-LTNKERKALRKMYLDEQSSDINDADIDWEDGQDAPKVTQTD 285  
QY 268 AQETKRG-----RIQTKKEV-----SIKTLKELVHKVTSPEWMM 305  
DB 286 SATSKTGTSLIWKSCATKVTSKKEVANPVQOPSKLYSAQNTLDALFNVCFTPEDMIK 345  
QY 306 QPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFLILE--KAETSKLTNFSLPDTRT 363  
DB 346 QSDXYLELSELPNGPKINTILLHNQVKTSTMTAFDCIKFNEEEDKFLATIKDM-- 403  
QY 364 CRIFAFHGNVYKCHAI CCVLNRQGGKRNALVPHGPASTGKSIIAQIAQAVNGVCYN 423  
DB 404 -----GLNEQYLLKVLCTILTQGGKRGCIWFGPGGTGLLASLICKATVNYGMVT 456  
QY 424 AANVPFPNDCTNKLIVWEAGNFGQVNOQFKAI CSQTIRIDOKGSKQIEPTPVM 483  
DB 457 TSNPNFPTWDCGNRIIWAEECGNLGNWVEDFKAITGGGVKVDTKNKPQSKGCVIV 515  
QY 484 TTENITVVRIGCEERPEHTQPIRDRMLNIHL-----THLHPGDFGLVDKNEWPMICA 536  
DB 516 TSNITITVTVGCVETWAHAELPKQRMKIRCMKTIINPKTIPTG-----MLKR 564  
QY 537 WLKXNGYQSTWASYCAKWKV-----DWSENWAEKVPKPTPINLLGSARSPP-TTPKSTP 590  
DB 565 WL-----NTWDRQPIQLSHEMPELYLETTGPN-----SSATSATKTTGNSQP 606  
QY 591 LSONYALTPLASDLELALPWPSTPNTPVAGTA 623  
DB 607 TTAETASVSTADCD-----TPKRGAS 628

## RESULT 15

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N;Alternate names: nonstructural protein NS1; NS1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S52209

R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52209

A;Molecule type: DNA

A;Residues: 1-626 &lt;ZAD&gt;

A;Cross-references: UNIPROT:Q83288; EMBL:X75093; NID:g609091; PIDN:CAA52983.1; PID:g60909

A;Experimental source: strain FM

C;Genetics:

A;Gene: NS1

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 12.6%; Score 459.5; DB 2; Length 626;  
Best Local Similarity 25.3%; Pred. No. 2.3e-25;  
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKKOVFI FDSLVKCKLPEVLNTKNIF-----PGDVNWFVQHEWGK 122  
DB 45 WETGTGNWHEHVNPLWPTLADKI-----KNIFQRMNQFNQDETFDFQLEEGS 93  
QY 123 DQGMCHCVLIGKDFDFOAQGKWWRRQLNVYNSRWLVTACNVOLTPAERIK---LREIAED 179  
DB 94 EY-IHLHAVCPGECRFSVLGRY-----MSQIKDSILRDVYEG 129  
QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPPRDGGVFLSSDSG 234  
DB 130 KQVKIPDWFSSITTKRGGQNKVTAA-----YTLHYLIPKKQ-----PELQWAFNTMPLF 179  
QY 235 WKTNFKEGERHLVSKLYTDDMRPETVEITVTTAQETKRGRIQTKKEVSIKTTLKEIVHK 294  
DB 180 TAAALCQKQKQELDLDFQSEMNNAVQEDQAAPL-----ISNRAKNYSNVLVDLIEM 235  
QY 295 RVTSPEWMMQPDSDYIEMMAOPGGENLLKNTLEICTLTARTKTAFLILEKAETSKLT 354  
DB 236 GITSEKQWLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDT 295  
QY 355 NPSLPDTRCIRPAPFGWYVVKCHAI CCVLNRQGGKRNALVPHGPASTGKSIIAQIAQ 414  
DB 296 K-----NRIYQILKLNYPQVYVGVLCGWVKREFNKRNAIMLYGPATTKTNIAEAIH 350  
QY 415 AVGNVCYNAANVPFPNDCTNKLIVWEAGNFGQVNOQFKAI CSQTIRIDOKGSKG 474  
DB 351 AVPFYGVVNTNENFPFNDQVDMGLIWEEGKMTNKVYESAKAILGGSVAVVDQCKGVS 410  
QY 475 QIEPTPVIITNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLHPGDFGLVDKNE---- 530  
DB 411 CIEPTPVIITNTDMCMIVDGNSTTWEHRIPLEERFQIVLSHKLGNFGKISKKEVKEF 470  
QY 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDSEN-----WAEPK----- 568  
DB 471 FKWANDNLVPVSEFVKVPTNEQTKLTE-----PVPERANEPSEPPPKIWAAPTRELEEI 524  
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPWPSTPNT 617  
DB 525 LRASPELFSVAPLP-----SSPDTSPKPKTKRGEYQVRCAMHSL-DNSMNVFECLEC 576  
QY 618 PVAGTATQNTGBAGSKAC-QDQQLSPWTWSEIIEEDLRACFGAEPPLKKDFSE 667  
DB 577 ERANPPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

Search completed: January 22, 2005, 03:42:03

Job time : 24.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:19:28 ; Search time 106.333 Seconds  
(without alignments)  
3636.223 Million cell updates/sec

Title: US-10-069-056-14

Perfect score: 3637

Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKKDFSEPLNLD 672

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3623	99.6	672	1	VNCS_MUMIV	P03134 murine minu
2	3623	99.6	721	2	Q84365	Q84365 murine minu
3	3554	97.7	721	1	VNCS_MUMIM	P07300 murine minu
4	3554	97.7	721	2	Q84363	Q84363 murine minu
5	3552	97.7	672	2	Q83429	Q83429 mouse parvo
6	3345	92.0	672	2	Q8JVL4	Q8JVL4 rat minute
7	3336	91.7	672	2	Q8JVL8	Q8JVL8 kilham rat
8	3335	91.7	672	2	P88899	P88899 kilham rat
9	3327	91.5	672	2	Q8JVL8	Q8JVL8 hamster par
10	3326	91.4	672	1	VNCS_PAVHH	P03133 hamster par
11	3323	91.4	672	2	Q8JVL6	Q8JVL6 rat minute
12	3322	91.3	668	1	VNCS_PAVL3	P36311 parvovirus
13	3303	90.8	665	2	O71159	O71159 kilham rat
14	2935.5	80.7	671	2	O71157	O71157 rat parvovi
15	2655	73.0	668	2	P89513	P89513 feline panl
16	2654	73.0	668	2	P89516	P89516 feline panl
17	2653	72.9	668	2	P89512	P89512 feline panl
18	2653	72.9	668	2	P89515	P89515 feline panl
19	2653	72.9	668	2	P90449	P90449 feline panl
20	2653	72.9	668	2	P90484	P90484 feline panl
21	2652	72.9	668	2	P89514	P89514 feline panl
22	2652	72.9	668	2	P90472	P90472 feline panl
23	2647	72.8	668	1	VNCS_MEVA	P27438 mink enteri
24	2644	72.7	668	1	VNCS_FPV19	P24842 feline panl
25	2642	72.6	668	1	VNCS_PAVCN	P12929 canine parv
26	2637	72.5	668	2	Q84393	Q84393 canine parv
27	2614	71.9	668	2	Q70M74	Q70M74 canine parv
28	2614	71.9	668	2	CAE47433	CAE47433 canine pa
29	2485	68.3	660	1	VNCS_PAVPN	P18547 porcine par
30	2476	68.1	662	1	VNCS_PAVPK	P52502 porcine par
31	2472	68.0	662	2	Q6RED5	Q6RED5 porcine par

32	2472	68.0	662	2	Q6TDD8	Q6TDD8 porcine par
33	2472	68.0	662	2	AAQ90279	AAQ90279 porcine p
34	2472	68.0	662	2	AA91039	AA91039 porcine p
35	2463	67.7	662	2	Q6PS60	Q6PS60 porcine par
36	2463	67.7	662	2	AA93262	AA93262 porcine p
37	1933	53.1	397	2	Q993M6	Q993M6 autonomous
38	1821.5	44.6	392	1	VNCS_FPV	P06431 feline panl
39	551	15.1	641	2	Q65017	Q65017 aleutian mi
40	545.5	15.0	641	2	Q65020	Q65020 aleutian mi
41	544	15.0	590	1	VNCS_ADVG	P24030 aleutian mi
42	543	14.9	620	2	Q96607	Q96607 aleutian mi
43	534.5	14.7	641	2	Q65023	Q65023 aleutian mi
44	460.5	12.7	610	2	Q6JL80	Q6JL80 bovine aden
45	460.5	12.7	610	2	AA26464	AA26464 bovine ad

#### ALIGNMENTS

RESULT 1  
VNCS\_MUMIV ID VNCS\_MUMIV STANDARD; PRT; 672 AA.  
AC P03134;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVBL1).  
GN Name=NS1;  
OS Murine minute virus (Murine parvovirus).  
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
OX NCBI\_TaxID=10794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83143341; PubMed=6298737;  
RA Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;  
RT "The complete DNA sequence of minute virus of mice, an autonomous  
RT parvovirus.";  
RL Nucleic Acids Res. 11:999-1018(1983).  
RN [2]  
RP HOMOLOGIZATION, AND INTERACTION WITH SYNCRIP.  
RX MEDLINE=99102562; PubMed=9847309;  
RA Harris C.E., Boden R.A., Astell C.R.;  
RT "A novel heterogeneous nuclear ribonucleoprotein-like protein  
RT interacts with NS1 of the minute virus of mice.";  
RL J. Virol. 73:72-80(1999).  
CC -I- FUNCTION: Seems necessary for viral DNA replication.  
CC -I- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP.  
CC -I- DOMAIN: The N-terminus (residues 1-275) possesses a negative effect  
CC on transactivation.  
CC -I- DOMAIN: The C-terminus (residues 543-672) possesses an activation  
CC domain.  
CC -I- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; J02275; AAA67109.1; -;  
EMBL; V01115; CAA24309.1; ALT\_INIT.  
PIR; A03696; UYFVIM.  
TRANSFAC; T02375; -;  
InterPro; IPR001257; Parvo NS1.  
Pfam; PF01057; Parvo NS1; 1.  
KW ATP-binding; DNA replication; Noncapsid protein;  
KW Nonstructural protein.  
FT DOMAIN 1 276 Interacts with SYNCRIP.  
FT NP\_BIND 399 406 ATP (Potential).  
SQ SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;

Query Match		99.6%; Score 3623; DB 1; Length 672;
Best Local Similarity		99.7%; Pred. No. 3.5e-259;
Matches	670; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1	MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPKNENVOLNGKDIGNYSYKKELODELKSLQ 60
DB	1	MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPKNENVOLNGKDIGNYSYKKELODELKSLQ 60
QY	61	RGATTWDSQSEDEMEWETTVDEMTKKQVIFPDSLVKKCLFEVLTAKNIIPGDDVNNFVQHEW 120
DB	61	RGATTWDSQSEDEMEWETTVDEMTKKQVIFPDSLVKKCLFEVLTAKNIIPGDDVNNFVQHEW 120
QY	121	GKQDQWHCHVLIGKDFSOAQGKWRRLNLYVWSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB	121	GKQDQWHCHVLIGKDFSOAQGKWRRLNLYVWSRWLVTAACNVQLTPAERIKLREIAEDN 180
QY	181	EWTLTYKHQKQDYTKCVLFGNMIAIYFLTKKISTSPRRDGGYFLSSDSGKNTNLF 240
DB	181	EWTLTYKHQKQDYTKCVLFGNMIAIYFLTKKISTSPRRDGGYFLSSDSGKNTNLF 240
QY	241	KEGERHLVSKLYTDDMRPEVETVTTAOTKKGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
DB	241	KEGERHLVSKLYTDDMRPEVETVTTAOTKKGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
QY	301	DWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB	301	DWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
QY	361	TRTCRIFAFHGNVYKVAICCVLNRQGGKRNVLVPHGPASTGKSIIAQAIQAQVGNVG 420
DB	361	TRTCRIFAFHGNVYKVAICCVLNRQGGKRNVLVPHGPASTGKSIIAQAIQAQVGNVG 420
QY	421	CYNAANVPFPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKGGSKQIEPTP 480
DB	421	CYNAANVPFPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKGGSKQIEPTP 480
QY	541	NGYSTWASCAKWKGVDPSENWAEKVPPTINLIGSARSPTTPKSTPLSQNYALTPL 600
DB	541	NGYSTWASCAKWKGVDPSENWAEKVPPTINLIGSARSPTTPKSTPLSQNYALTPL 600
QY	601	ASLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEP 660
DB	601	ASLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEP 660
QY	661	LKXDFSEPLNLD 672
DB	661	LKXDFSEPLNLD 672
RESULT 2		
ID	Q84365	PRELIMINARY; PRT; 721 AA.
AC	Q84365;	
DT	01-NOV-1996	(TrEMBLrel. 01, Created)
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)
DE	Nonstructural protein.	
GN	Name=NS1;	
OS	Murine minute virus (Murine parvovirus).	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	
OX	NCBI_TaxID=10794;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=MVM;	
RX	MEDLINE=83143341; PubMed=6298737;	
RA	Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;	
RT	"The complete DNA sequence of minute virus of mice, an autonomous	
RT	parvovirus."	
RL	Nucleic Acids Res. 11:999-1018(1983).	
[2]		
SEQUENCE FROM N.A.		
STRAIN=MVM;		
MEDLINE=86115415; PubMed=3502703;		
Astell C.R., Gardiner E.M., Tattersall P.;		
"DNA sequence of the lymphotropic variant of minute virus of mice,		
MVM(i), and comparison with the DNA sequence of the fibrotropic		
RT prototype strain."		
J. Virol. 57:656-669(1986).		
[3]		
SEQUENCE FROM N.A.		
STRAIN=MVM;		
MEDLINE=87061199; PubMed=3783817;		
Morgan W.R., Ward D.C.;		
"Three splicing patterns are used to excise the small intron common to		
all minute virus of mice RNAs."		
J. Virol. 60:1170-1174(1986).		
EMBL; J02275; AAA67108.1; -		
GO; GO:0019012; C:virion; IEA.		
GO; GO:0019079; P:viral genome replication; IEA.		
InterPro; IPR001257; Parvo.NS1.		
Pfam; PF01057; Parvo.NS1; I.		
Nonstructural protein.		
SEQUENCE 721 AA; 81896 MW; 18391758E42F0DCF CRC64;		
Query Match		99.6%; Score 3623; DB 2; Length 721;
Best Local Similarity		99.7%; Pred. No. 3.8e-259;
Matches	670; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1	MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPKNENVOLNGKDIGNYSYKKELODELKSLQ 60
DB	50	MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPKNENVOLNGKDIGNYSYKKELODELKSLQ 109
QY	61	RGATTWDSQSEDEMEWETTVDEMTKKQVIFPDSLVKKCLFEVLTAKNIIPGDDVNNFVQHEW 120
DB	110	RGATTWDSQSEDEMEWETTVDEMTKKQVIFPDSLVKKCLFEVLTAKNIIPGDDVNNFVQHEW 169
QY	121	GKQDQWHCHVLIGKDFSOAQGKWRRLNLYVWSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB	170	GKQDQWHCHVLIGKDFSOAQGKWRRLNLYVWSRWLVTAACNVQLTPAERIKLREIAEDN 229
QY	181	EWTLTYKHQKQDYTKCVLFGNMIAIYFLTKKISTSPRRDGGYFLSSDSGKNTNLF 240
DB	230	EWTLTYKHQKQDYTKCVLFGNMIAIYFLTKKISTSPRRDGGYFLSSDSGKNTNLF 289
QY	241	KEGERHLVSKLYTDDMRPEVETVTTAOTKKGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
DB	290	KEGERHLVSKLYTDDMRPEVETVTTAOTKKGRIQTKKEVSIKTTLLKELVHKRVTSPE 349
QY	301	DWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB	350	DWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 409
QY	361	TRTCRIFAFHGNVYKVAICCVLNRQGGKRNVLVPHGPASTGKSIIAQAIQAQVGNVG 420
DB	410	TRTCRIFAFHGNVYKVAICCVLNRQGGKRNVLVPHGPASTGKSIIAQAIQAQVGNVG 469
QY	421	CYNAANVPFPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKGGSKQIEPTP 480
DB	470	CYNAANVPFPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKGGSKQIEPTP 529
QY	481	VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540
DB	530	VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 589
QY	541	NGYSTWASCAKWKGVDPSENWAEKVPPTINLIGSARSPTTPKSTPLSQNYALTPL 600
DB	590	NGYSTWASCAKWKGVDPSENWAEKVPPTINLIGSARSPTTPKSTPLSQNYALTPL 649
QY	601	ASLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEP 660
DB	650	ASLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEDLRACFGAEP 709



```
QY 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 3
VNC5 MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Name=NS1;
OS Murine minute virus (strain MVM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP (By
CC similarity).
CC -1- DOMAIN: The N-terminus (residues 1-275) possesses a negative effect
CC on transactivation (By similarity).
CC -1- DOMAIN: The C-terminus (residues 543-672) possesses an activation
CC domain (By similarity).
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02481; -; NOT ANNOTATED_CDS.
DR EMBL; M12032; AAA69567.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT DOMAIN 1 276 Interacts with SYNCRIP (By similarity).
FT NP BIND 399 406 ATP (Potential).
FT CONFLICT 597 597 I -> L (in Ref. 2).
SQ SEQUENCE 672 AA; 25F025FE32B84DF0 CRC64;

Query Match 97.7%; Score 3554; DB 1; Length 672;
Best Local Similarity 97.5%; Pred. No. 4.4e-254;
Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSPFNENVQLNGDKIGNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKEKSNQEVFSPFNENVQLNGDKIGNSYKKELODELKSLQ 60.
QY 61 RGAETTWQSQEDMEWETTVDEMTKQVFIPIFDSLKKCLFEVLATNKNIFPGDVNWFVQHEW 120
Db 61 RGAETTWQSQEDMEWETTVDEMTKQVFIPIFDSLKKCLFEVLATNKNIFPGDVNWFVQHEW 120
QY 121 GKQDQGWCHVLIIGKDFSOQGWRRQLNVLVYTSRWLVLTACNVQLTPAERIKLREIAEDN 180

RESULT 4
Q84363 PRELIMINARY; PRT; 721 AA.
AC Q84363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NS1;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
DR EMBL; M12032; AAA69566.1; -.
DR PIR; A23008; UYPVIM.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein.
SQ SEQUENCE 721 AA; 81862 MW; 9FD29C327C7F4BBF CRC64;

Query Match 97.7%; Score 3554; DB 2; Length 721;
Best Local Similarity 97.5%; Pred. No. 4.9e-254;
Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGNWSYKKEQLQDELKSLQ 60
Db 50 MAGNAYSDEVLTGTTNMLKEKSNQEVFSFVKETDVLQNGKDIGNWSYKKEQLQDELKSLQ 109
QY 61 RGAETTTWQSEDMEWETTVDKQVFIYDLSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW 120
Db 110 RGAETTTWQSEDMEWETTVDKQVFIYDLSLVKKCLFEVLNTKNIAPADVTWTFVQHEW 169
QY 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLNVYMSRWLVACNVQLTPAERIKLREIAEDN 180
Db 170 GKQDQGHCHVLIGGKDFSOAQGKWRRLNVYMSRWLVACNVQLTPAERIKLREIAEDS 229
QY 181 EWTLLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
Db 230 EWTLLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 289
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIQTKKEVSIKTTLKELVHKRVTSPE 300
Db 290 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIQTKKEVSIKTTLKELVHKRVTSPE 349
QY 301 DWMMPQDSYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360
Db 350 DWMMPQDSYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 409
QY 361 TRTCRIFAFHGMNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 420
Db 410 TRTCRIFAFHGMNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 469
QY 421 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDOKGKSKQIEPTP 480
Db 470 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDOKGKSKQIEPTP 529
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 589
QY 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
Db 590 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 649
QY 601 ASDLEDLALPEWSTPNTPVAGTAETONTGEAGSKACODGQLSPTWSEIEDLRACFGAEP 660
Db 650 ASDLEDLALPEWSTPNTPVAGTAETONTGEAGSKACODGQLSPTWSEIEDLRACFGAEP 709
QY 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 5
Q83429 ID Q83429 PRELIMINARY; PRT; 672 AA.
AC Q83429;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NSI;
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP MEDLINE=94365951; PubMed=8083985;
RX Ball-Goodrich L.J., Johnson E.;
RT "Molecular characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486(1994).
DR EMBL; U12469; AAA61405.1;
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:virion genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; I.
KW Nonstructural protein.
```

```
SQ SEQUENCE 672 AA; 76112 MW; 31C6365276727363 CRC64;
Query Match 97.7%; Score 3552; DB 2; Length 672;
Best Local Similarity 97.3%; Pred. No. 6.2e-254;
Matches 654; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGNWSYKKEQLQDELKSLQ 60
Db 1 MAGNAYSDEVLTGTTNMLKEKSNQEVFSFVKETDVLQNGKDIGNWSYKKEQLQDELKSLQ 60
QY 61 RGAETTTWQSEDMEWETTVDKQVFIYDLSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW 120
Db 61 RGAETTTWQSEDMEWETTVDKQVFIYDLSLVKKCLFEVLNTKNIAPADVTWTFVQHEW 120
QY 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLNVYMSRWLVACNVQLTPAERIKLREIAEDN 180
Db 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLNVYMSRWLVACNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIQTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIQTKKEVSIKTTLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360
Db 301 DWMMPQDSYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360
QY 361 TRTCRIFAFHGMNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 420
Db 361 TRTCRIFAFHGMNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 420
QY 421 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDOKGKSKQIEPTP 480
Db 421 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDOKGKSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
Db 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPEWSTPNTPVAGTAETONTGEAGSKACODGQLSPTWSEIEDLRACFGAEP 660
Db 601 ASDLEDLALPEWSTPNTPVAGTAETONTGEAGSKACODGQLSPTWSEIEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 6
Q8JUV14 ID Q8JUV14 PRELIMINARY; PRT; 672 AA.
AC Q8JUV14;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NSI;
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP MEDLINE=22120170; PubMed=12124471;
RX Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat parvoviruses.";
```

```
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venemo M., Pintel D., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF321284; AM93279.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSL.
DR Pfam; PF01057; Parvo_NSL; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 92.0%; Score 3345; DB 2; Length 672;
Best Local Similarity 91.7%; Pred. No. 1.3e-238;
Matches 616; Conservative 22; Mismatches 34; Indels 0; Gaps 0;

QY 1 MAGNAYSDVLGATNMLKEKSNOEVSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDVLGATNMLKDKSSQEVFSFVKNNVQLNGKDIGNSYRKELODELKSLQ 60
QY 61 RGAETTWQSDSEMEWETTVDEMTHKQVFIKQVIFDLSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW 120
DB 61 RGAETTWQSDSEMEWESAVDDMTKQVFIKQVIFDLSLVKKCLFEVLNTKNIAPSDVTFVQHEW 120
QY 121 GKQGGHCHVLIGKDFSOAQGWRRQLNYYWSRMLVTACNVOLTPAERIKLREIAEDN 180
DB 121 GKQGGHCHVLIGKDFSOAQGWRRQLNYYWSRMLVTACNVOLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHKOTKDYKCVLFGNMIAYVFLTKKISTSPPRDGGYFSSDSGKTNFL 240
DB 181 EWTLLTYKHKOTKDYKCVLFGNMIAYVFLSKKICTSPPRDGGYFSSDSGKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPEVETVTTAQTKEGRIOTKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMKEPEVETVTTAQAEGRIOTRKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMMPQDSYIENMAOPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLNFSLPD 360
DB 301 DNMMPQDSYIENMAOPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLNFSLPD 360
QY 361 TRTCRIFAEHGWNYKCHAIACVLRNQGKRNALVPHGPASTGKSIIAQAIAGVGNVG 420
DB 361 TRTCRIFAEHGWNYKCHAIACVLRNQGKRNALVPHGPASTGKSIIAQAIAGVGNVG 420
QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQIRIDQKGGSKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVDPWSENWABPKVPTPINLLGSARSPTTPKSTPLSONYALTPL 600
DB 541 NGYQSTWASYCAKWKVDPWSENWABPKVPTPINLLGSARSPTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACODGQSLPTWSEIEDLACFAGEP 660
DB 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACODGQSLPTWSEIEDLACFAGEP 660
QY 661 LKDFSEPLNLD 672
DB 661 LKDFSEPLNLD 672

RESULT 7
Q8JV28 PRELIMINARY; PRT; 672 AA.
ID Q8JV28
AC Q8JV28
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Nonstructural protein 1.
GN Name=ns1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22124471;
RA Wan C.H., Soderlund-Venemo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses.";
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Soderlund-Venemo M., Pintel D., Riley L.K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF321230; AM93272.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSL.
DR Pfam; PF01057; Parvo_NSL; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75987 MW; 22B4611C20CDB6E9 CRC64;

Query Match 91.7%; Score 3336; DB 2; Length 672;
Best Local Similarity 91.5%; Pred. No. 6e-238;
Matches 615; Conservative 22; Mismatches 35; Indels 0; Gaps 0;

QY 1 MAGNAYSDVLGATNMLKEKSNOEVSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDVLGATNMLKDKSSQEVFSFVKNNVQLNGKDIGNSYRKELODELKSLQ 60
QY 61 RGAETTWQSDSEMEWETTVDEMTHKQVFIKQVIFDLSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW 120
DB 61 RGAETTWQSDSEMEWESAVDDMTKQVFIKQVIFDLSLVKKCLFEVLNTKNIAPSDVTFVQHEW 120
QY 121 GKQGGHCHVLIGKDFSOAQGWRRQLNYYWSRMLVTACNVOLTPAERIKLREIAEDN 180
DB 121 GKQGGHCHVLIGKDFSOAQGWRRQLNYYWSRMLVTACNVOLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHKOTKDYKCVLFGNMIAYVFLTKKISTSPPRDGGYFSSDSGKTNFL 240
DB 181 EWTLLTYKHKOTKDYKCVLFGNMIAYVFLSKKICTSPPRDGGYFSSDSGKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPEVETVTTAQTKEGRIOTKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMKEPEVETVTTAQAEGRIOTRKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMMPQDSYIENMAOPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLNFSLPD 360
DB 301 DNMMPQDSYIENMAOPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLNFSLPD 360
QY 361 TRTCRIFAEHGWNYKCHAIACVLRNQGKRNALVPHGPASTGKSIIAQAIAGVGNVG 420
DB 361 TRTCRIFAEHGWNYKCHAIACVLRNQGKRNALVPHGPASTGKSIIAQAIAGVGNVG 420
QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQIRIDQKGGSKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVDPWSENWABPKVPTPINLLGSARSPTTPKSTPLSONYALTPL 600
DB 541 NGYQSTWASYCAKWKVDPWSENWABPKVPTPINLLGSARSPTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACODGQSLPTWSEIEDLACFAGEP 660
DB 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACODGQSLPTWSEIEDLACFAGEP 660
QY 661 LKDFSEPLNLD 672
```



Db	181	EWTLLTYKHKHTKDYTKCVLFGNMIAYYFLSKKKICTSPRDCGGYFLSSDSGKTNFL	240
Qy	241	KEGERHLVSKLYTDDMEPEVETVTVTTAQETKRGRIOTKKEVSIKTTLKELVHKVTSPE	300
Db	241	KEGERHLVSKLYTDDMEPEVETVTVTTAQEAKRGRIQTRBEVSIKTTLKELVHKVTSPE	300
Qy	301	DWMQPDSDYIEMMAQPGGENLLKNTLLEICTPLTLARTKTAFDLILEKAETSKLTNPSLDP	360
Db	301	DWMQPDSDYIEMMAQPGGENLLKNTLLEICTPLTLARTKTAFDLILEKAETSKLANFMSAS	360
Qy	361	TRTCRIFAFAHGNVYKVKCHALCCVLNRQGGKRNVLPHGPASTGKSIITAAIAQAVGNVG	420
Db	361	TRTCRIFAFAHGNVYKVKCHALCCVLNRQGGKRNVLPHGPASTGKSIITAAIAQAVGNVG	420
Qy	421	CYNAANVNFPPDCTNKNLIWVEEAGNFGQOVNQFKALCSQOTIRIDOKGKSGKQIEPTP	480
Db	421	CYNAANVNFPPDCTNKNLIWVEEAGNFGQOVNQFKALCSQOTIRIDOKGKSGKQIEPTP	480
Qy	481	VIMTTNENITVVRIGCSERPEHTOPIRDRMLNIHLTHHLPGDGLVLDKNEWPMICAWLVK	540
Db	481	VIMTTNENITVVRIGCSERPEHTOPIRDRMLNIHLTRTLPDGLVLDKHEWPLICAWLVK	540
Qy	541	NGYQSTWASCAKWKVPPDRSENWAEKVPPTPINLGSARSPTTPKSTPLSONYALTPL	600
Db	541	NGYQSTWACCAKWKVPPDSEDAEPKLDTPINSLGSMRSPSLTPRSTPLSQNYALTPL	600
Qy	601	ASDLEDLALPSPWSTNPTPVAGTATONTGEAGSKACODGQLSPTWSEIEEDLRACFGAEP	660
Db	601	ASDLADLALPSPWSTNPTPVAGTAAQNTGEAGSTACQGAQSPTWSEIEADLRACFSQEQ	660
Qy	661	LKXDFSEPLNLD 672	
Db	661	LESDFNEBELTD 672	

RESULT 11

Q8JV16

ID	Q8JV16	PRELIMINARY;	PRT;	672 AA.
AC	Q8JV16;			
DT	01-OCT-2002 (TrEMBLrel. 22, Created)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)			
DE	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Nonstructural protein 1.			
GN	Name=NS1;			
OS	Rat minute virus 1b.			
OX	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.			
NCBI_TaxID=172386;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22120170; PubMed=12124471;			
RA	Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;			
RT	"Molecular characterization of three newly recognized rat			
RV	parvoviruses."			
RL	J. Gen. Virol. 83:2075-2083(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Wan C.-H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF332883; AAM93277.1; -.			
DR	GO; GO:0019012; C:virion; IEA.			
DR	GO; GO:0019079; P:Viral genome replication; IEA.			
DR	InterPro; IPR001257; Parvo NS1.			
DR	Pfam; PF01057; Parvo NS1; I.			
KW	Nonstructural protein_1			
SQL	SEQUENCE 672 AA; 76201 MW; C2FLA71F6EF449A6 CRC64;			

Query Match 91.4%; Score 3323; DB 2; Length 672;

Best Local Similarity 90.9%; Pred. No. 5.5e-237;

Matches 611; Conservative 24; Mismatches 37; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESKNQBVFSFVKENVQLNGKQIGNWSYKKEIQEDELKSLQ 60



```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-Mass;
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup.";
RL J. Virol. 72:3289-3299(1998).
DR EMBL; AF036711; AAC40695.1; -.
DR GO; GO:0019012; C:viral genome replication; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; I.
KW Nonstructural protein.
FT NON TER 1
SQ SEQUENCE 665 AA; 75375 MW; 778B29043417E409 CRC64;

Query Match 90.8%; Score 3303; DB 2; Length 665;
Best Local Similarity 91.4%; Pred. No. 1.6e-235;
Matches 608; Conservative 22; Mismatches 35; Indels 0; Gaps 0;

Qy 8 DEVLGATNMLKESNOEVSFVKFNENVLNGKDIGNSYKKELODELKSLQGAETTW 67
Db 1 DEVLGATNMLKESNOEVSFVKFNENVLNGKDIGNSYKKELODELKSLQGAETTW 60

Qy 68 DQSEDMEWETTDVEMTKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFQHEWKGQGH 127
Db 61 DQSEDMEWESAVDMTKKQVIFDLSLVKKCLFEVLNTKNIAPSDVTFVQHEWKGQGH 120

Qy 128 CHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERI KLEIAEDNEWTLLT 187
Db 121 CHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERI KLEIAEDSEWTLT 180

Qy 188 YKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSDSGWKNFLKEGERHL 247
Db 181 YKHHTKDYTKCVLFGNMIAYFLSKKICTSPRDGGYFLSDSGWKNFLKEGERHL 240

Qy 248 VSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPEDEMMQ 307
Db 241 VSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPEDEMMQ 300

Qy 308 DSVIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPSLPDTRCRIF 367
Db 301 DSVIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANPFMASTRCRIF 360

Qy 368 AFHGMVYKCHAIACVLRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNCYNAANV 427
Db 361 AFHGMVYKCHAIACVLRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNCYNAANV 420

Qy 428 NFPFNDCTNKLWEEAGNFGQVNOFKAIICSGQTIRIDQKGSKQIEPTVIMTNE 487
Db 421 NFPFNDCTNKLWEEAGNFGQVNOFKAIICSGQTIRIDQKGSKQIEPTVIMTNE 480

Qy 488 NITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVKNGYOSTM 547
Db 481 NITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVKNGYOSTM 540

Qy 548 ASYCAKWKVPDSEWAEKVPPTPINLIGSARSPTTPKSTPLSONYALTPLASLEDL 607
Db 541 ASYCAKWKVPDSEWAEKVPPTPINLIGSARSPTTPKSTPLSONYALTPLASLEDL 600

Qy 608 ALPEWSTPNTPVAGTAETQNTGAGSKACODGQLSPWSEIEDLRACFGAEPLKDFSE 667
Db 601 ALPEWSTPNTPVAGTAETQNTGAGSKACODGQLSPWSEIEDLRACFGAEPLKDFSE 660

Qy 668 PLNLD 672
```

```

Db 661 EBLTD 665

RESULT 14
O71157 PRELIMINARY; PRT; 671 AA.
AC O71157;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NS1;
OS Rat parvovirus 1a.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=74581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup.";
RL J. Virol. 72:3289-3299(1998).
DR EMBL; AF036710; AAC40693.1; -.
DR GO; GO:0019012; C:viral genome replication; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; I.
KW Nonstructural protein.
SQ SEQUENCE 671 AA; 75752 MW; 98CB39A39298D4DE CRC64;

Query Match 80.7%; Score 2935.5; DB 2; Length 671;
Best Local Similarity 81.1%; Pred. No. 2.6e-208;
Matches 545; Conservative 42; Mismatches 84; Indels 1; Gaps 1;

Qy 1 MAGNAYSDEVLGATNMLKESNOEVSFVKFNENVLNGKDIGNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNMLKESNOEVSFVKFNENVLNGKDIGNSYKKELODELKSLQ 60

Qy 61 RGAETTWDSQEDMEWETTDVEMTKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFQHEW 120
Db 61 RGAETTWDSQEDMEWETTDVEMTKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFQHEW 120

Qy 121 GKDGQWHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERI KLEIAEDN 180
Db 121 GKDGQWHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERI KLEIAEDN 180

Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSDSGWKNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSDSGWKNFL 240

Qy 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300

Qy 301 DNMAMPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPSLPD 360
Db 301 DNMAMPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNPSLPD 360

Qy 361 TRTCRIFAFHGMVYKCHAIACVLRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNCV 420
Db 361 TRTCRIFAFHGMVYKCHAIACVLRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNCV 420

Qy 421 CYNAAVNFPENDCTNKLWEEAGNFGQVNOFKAIICSGQTIRIDQKGSKQIEPTP 480
Db 421 CYNAAVNFPENDCTNKLWEEAGNFGQVNOFKAIICSGQTIRIDQKGSKQIEPTP 480

Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540

Qy 541 NGYQSTMASYCAKWKVPDSEWAEKVPPTPINLIGSARSPTTPKSTPLSONYALTPL 600
```

```
Db 541 NGYOSTWASYSHHWGKVPDWSENWAEPAWTPINSIGSARSTQSTATSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGAGSKACODGQLSPWSEIEDLRACIFCAEP 660
Db 601 ASDLADLALPWSPTNPVAFVPTGTSNT-NTGGRNSQTARASPTWSEIADLRACFSOEQ 659
Qy 661 LKKDFSEPLNLD 672
Db 660 LESDFNEELTLD 671

RESULT 15
P89513 ID P89513 PRELIMINARY; PRT; 668 AA.
AC P89513;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RA Horiuchi M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB000053; BAA19014.1; -.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:000166; F:nucleotide binding; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR001257; Parvo NS1.
DR Pfam: PF01057; Parvo NS1; I.
DR SMART: SMO0382; AAA; 1.
KW Nonstructural protein.
SQ SEQUENCE 668 AA; 76743 MW; 84A210104D9D5788 CRC64;
```

```
Query Match 73.0%; Score 2655; DB 2; Length 668;
Best Local Similarity 73.4%; Pred. No. 1.5e-187;
Matches 493; Conservative 70; Mismatches 103; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPNENVLQNGKDIGWNSYKKELOEDELKSLQ 60
Db 1 MSGNQYTEEVWEGVNLKKAENAEAFSEFKCDNVQNGKDVRRWNTKPIQNEELTSLI 60

Qy 61 RGAETTWDS--EDMEHTTVDENTKKQVIFDSLKKCLFEVLNTKNIFFPGDVNHFVQH 118
Db 61 RGAQTAMDQTEEBEMDWESEVDSLAKKQVTFDALIKKCLFEVFSKNIEPNECVWFIOH 120

Qy 119 EWGKDQGHCHVLIGGKDFGQAGKWRRLQVNVYWRWLVTACNVOLTPAERIKLREIAE 178
Db 121 EWGKDQGHCHVLLSHSKNLQATGKWLRRQNNYNSRWLVLCVNLTPTEKIKLREIAE 180

Qy 179 DNEWVTLTYKHQTKDYKCVLFQGMIAFYFLTKKISTSPRDGGYFLSSDSGWKTN 238
Db 181 DSEWVTLTYRHQTKDYKCVLVHFGNMIAFYFLTKKKI-VHMTKESGYFLSTDGSKFN 239

Qy 239 FLKEGERHLSKLYTDDMRPETVTTTAQETTKRGRIQTKKEVSIKTKILKELVHKRVTS 298
Db 240 FMKHQDRHTVSTLYTEQKMPETVTTTAQETTKRGRIQTKKEVSIKTLRDLVSKRVTS 299

Qy 299 PEDWMMMQPDSYIEMWAPQGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSL 358
Db 300 PEDWMMMQPDSYIEMWAPQGGENLLKNTLEICTLTARTKTAFLDILEKADNTKLTNFDL 359

Qy 359 PDTTRCFIFAHGNNYKVKCHAI CCVLNRQGGKRNVLFFHGPASTGKSIIAQIAQAVGN 418
Db 360 ANSRTQIFRMHGNWLVKCHAIACVLNRQGGKRNVLFFHGPASTGKSIIAQIAQAVGN 419

Qy 419 VGCYNAANVNFENDCTNKNLIWVEEAGNFGQVNOFKALCSGOTIRIDQKGGSKOIEP 478
Db 420 VGCYNAANVNFENDCTNKNLIWIEEAGNFGQVNOFKALCSGOTIRIDQKGGSKOIEP 479
```

```
Qy 479 TPVIMTTNENITVVRIGCEBERPEHTOPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWL 538
Db 480 TPVIMTTNENITVVRIGCEBERPEHTOPIRDRMLNIKLKCKLPGDFGLVDKEWPLICAWL 539
Qy 539 VKNGYOSTWASYCAKWKGVPDWSENWAEPKVPTPINLGSARSFPTTPKSTPLSQNYALT 598
Db 540 VKHGYOSTWANYTHHWGKVPEDENWAEPKIQEGINSPG-CKOLETQAAASNPOSQDHVLT 598
Qy 599 PLASDLEDLALPWSPTNPVAGTAETQNTGAGSKACODGQLSPWSEIEDLRACIFGA 658
Db 599 PLTPDVVDLALPWSPTNPVAGTAETQNTGAGSKACODGQLSPWSEIEDLRACIFGA 658
Qy 659 EPLKKDFSEPLN 670
Db 657 EQLEEDFRDDLD 668
```

Search completed: January 22, 2005, 03:40:47  
Job time : 108.333 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds  
(without alignments)  
11177.792 Million cell updates/sec

Title: US-10-069-056-16  
Perfect score: 2019  
Sequence: 1 atggctggaatgcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	GenEmbl.*			
	1: gb_ba.*	2: gb_hcg.*	3: gb_in.*	4: gb_cm.*
	5: gb_ov.*	6: gb_pat.*	7: gb_ph.*	8: gb_pl.*
	9: gb_pr.*	10: gb_ro.*	11: gb_sts.*	12: gb_sy.*
	13: gb_un.*	14: gb_vi.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2019	100.0	2019	6	AX137751 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	V01115 Minute viru
4	2017.4	99.9	5149	14	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137739 Sequence
6	2015.8	99.8	2019	6	AX137743 Sequence
7	2015.8	99.8	2019	6	AX137747 Sequence
8	1905.4	94.4	5085	14	M12032 Minute viru
9	1902.2	94.2	4764	14	U34253 Mouse parvo
10	1899	94.1	5087	14	X02481 Mouse parvo
11	1899	94.1	5144	14	U12469 Mouse parvo
12	1873	92.8	4764	14	U34254 Mouse parvo
13	1862.2	92.2	4773	14	U34255 Hamster par
14	1855.8	91.9	4761	14	U34256 Mice minute
15	1771	87.7	5121	6	CQ786765 Sequence
16	1662.2	82.3	4904	14	AF321230 Kilham ra
17	1659	82.2	4795	14	AF332884 Rat minut
18	1657.4	82.1	4927	14	U79033 Kilham rat
19	1656.6	82.1	5135	14	M81888 Parvovirus

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARH1	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	52.0	5124	14	PVPFVC	M8246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCY1A	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAPVNS1	X58115 Feline panl
40	1043.6	51.7	5124	14	PVCPVC	M8245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVMPFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137751	Sequence 16 from Patent EP1077260.	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137751					
DEFINITION	AX137751					
ACCESSION	AX137751.1	GI:14273925				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
CDS						
ORIGIN						
Query Match						
Score						
DB						
Length						

AX137751 Sequence 16 from Patent EP1077260.  
AX137751  
AX137751.1 GI:14273925  
Mice minute virus  
Mice minute virus  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
Nuesch, J. and Rommelaere, J.  
Parvovirus ns1 variants  
Patent: EP 1077260-A 16 21-FEB-2001;  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts  
(DE)  
Location/Qualifiers  
1..2019  
/organism="Mice minute virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10794"  
/note="Parvovirus NS1 variant"  
1..2019  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC39997.1"  
/db\_xref="GI:14273926"  
/translation="MAGNAYSDEVLGATNWLKESNQEVFSFVKHNVQLNGKIDIGW  
NSYKELQDELKSLQGAETTWDOSEDMEWETTVDEMTKQVFI FDSLVKKCLFEVL  
QTKKEVSIKTKLKVHRTSPEDMMQPDYSIEMMAQPGENLLKNTLEICLTLL  
ARTKAFDLILEKATSKLTNFSLPDTRTCRIFAFHGMVYKVCIAICVLEAGNFGQ  
NTVLFHGPASTGKSIIAQIAOAVGNVCYNAANVNFNFNDCTNKLIVVEAGNFGQ  
QVNPQKACSGAIRIDQKSKQIEPTVPMTNENITVVRIGCEEPHTQPIRD  
RMLNHLHTLPDGLFGLNWPIMCWLKXNGVQSTWASCAKMGKVPDSENWAE  
PKVPTPINLGSARSPFTPKSPPLSQNTALTPGLASDLEDLALSPWSTNTVPVGTAE  
TQNTGEAGSKACQDQQLSPTWSEIEDLACFGAEPLKDKFSEPLND"

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA	60				
Db	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA	60				
Qy	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAATA	120				
Db	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAATA	120				
Qy	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA	180				
Db	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA	180				
Qy	181	CGAGGAGCGGAATCTCTGGGACCAAGCAGGACATGAATGGGNAACCAACAGTGGAT	240				
Db	181	CGAGGAGCGGAATCTCTGGGACCAAGCAGGACATGAATGGGNAACCAACAGTGGAT	240				
Qy	241	GAAATGACCAAAAGCAAGTATTCTATTTTGAATCTTTGGTTAAATAAGTTTATTGAA	300				
Db	241	GAAATGACCAAAAGCAAGTATTCTATTTTGAATCTTTGGTTAAATAAGTTTATTGAA	300				
Qy	301	GTGCTTAACCAAGAATATATTTCCTGGTGAATTTAAATGGTTTGGCAACATGAATGG	360				
Db	301	GTGCTTAACCAAGAATATATTTCCTGGTGAATTTAAATGGTTTGGCAACATGAATGG	360				
Qy	361	GGAAAGACCAAGCTGGCACTGCCATGTACTTAATGGAGAAAGCACTTTAGTCAAGCT	420				
Db	361	GGAAAGACCAAGCTGGCACTGCCATGTACTTAATGGAGAAAGCACTTTAGTCAAGCT	420				
Qy	421	CAAGGGAATGTGGGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480				
Db	421	CAAGGGAATGTGGGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480				
Qy	481	TGTAATGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAAT	540				
Db	481	TGTAATGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAAT	540				
Qy	541	GAGTGGTGTACTTACTTACTTATAGCATAAGCAACCAAAAGCACTATACCAAGTGT	600				
Db	541	GAGTGGTGTACTTACTTACTTATAGCATAAGCAACCAAAAGCACTATACCAAGTGT	600				
Qy	601	GTTCCTTTTGGAAACATGATGTCTTACTATTTTTTAACTAAAAAGAAATAAGCACTAGT	660				
Db	601	GTTCCTTTTGGAAACATGATGTCTTACTATTTTTTAACTAAAAAGAAATAAGCACTAGT	660				
Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720				
Db	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720				
Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGCCAGAAACG	780				
Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGCCAGAAACG	780				
Qy	781	GTTGAAACCAACAGTAACCACTGGCGAGAAACTAAGCGCGGAGAAATCAAACTAAAAA	840				
Db	781	GTTGAAACCAACAGTAACCACTGGCGAGAAACTAAGCGCGGAGAAATCAAACTAAAAA	840				
Qy	841	GAGTTTCTATTAAACTACATTAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	900				
Db	841	GAGTTTCTATTAAACTACATTAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	900				
Qy	901	GACTGGATGATGAGCAGACAGTTACATTTGAATGATGCTCAACACAGGTGGAGAA	960				
Db	901	GACTGGATGATGAGCAGACAGTTACATTTGAATGATGCTCAACACAGGTGGAGAA	960				
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA	1020				
Db	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA	1020				
Qy	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACGAGAAATTAACCAACTTTTCACTGCCTGAC	1080				

Db	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACGAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
Qy	1081	ACAAGAACCTGCAGAAATTTTGGCTTTTTCATGCTGGCAACTATGCTAAAGTTTCCCATGCT	1140
Db	1081	ACAAGAACCTGCAGAAATTTTGGCTTTTTCATGCTGGCAACTATGCTAAAGTTTCCCATGCT	1140
Qy	1141	ATTGCTGTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATCTGTTTTATTTCATGGAACA	1200
Db	1141	ATTGCTGTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATCTGTTTTATTTCATGGAACA	1200
Qy	1201	GCAGACACAGGCAAAATCTATTATTGCAAGCCATAGCAAGAGCTTGGCAATGTTGGT	1260
Db	1201	GCAGACACAGGCAAAATCTATTATTGCAAGCCATAGCAAGAGCTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATGCGAGCAATGTAACTTTCCATTTTAATGACTGTACCAACAGAACTTGATT	1320
Db	1261	TGCTATAATGCGAGCAATGTAACTTTCCATTTTAATGACTGTACCAACAGAACTTGATT	1320
Qy	1321	TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTTCCTCT	1380
Db	1321	TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTTCCTCT	1380
Qy	1381	GGTCAAGCTATTTCGCATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA	1440
Db	1381	GGTCAAGCTATTTCGCATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA	1440
Qy	1441	GTCAATCATGACCAACCAATGAGAACATTACAGTGGTCAAGTGGTGGGCAAGAAAGACCA	1500
Db	1441	GTCAATCATGACCAACCAATGAGAACATTACAGTGGTCAAGTGGTGGGCAAGAAAGACCA	1500
Qy	1501	GAAACACATCAACCAATCGAGACAGATGCTTTAACTCATCTAAACATACCTTTGCCT	1560
Db	1501	GAAACACATCAACCAATCGAGACAGATGCTTTAACTCATCTAAACATACCTTTGCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTCCTGCTGTTGTTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTCCTGCTGTTGTTAAAG	1620
Qy	1621	AATGGTTTCAATCTACCATGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
Db	1621	AATGGTTTCAATCTACCATGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
Qy	1681	TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGTTTGGCACGC	1740
Db	1681	TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGTTTGGCACGC	1740
Qy	1741	TCACCAATTCAGACACCGAAAGTAGCTCTCTCAGCCAGAACTATGCACTTCCACTT	1800
Db	1741	TCACCAATTCAGACACCGAAAGTAGCTCTCTCAGCCAGAACTATGCACTTCCACTT	1800
Qy	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCTCTGTTGG	1860
Db	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCTCTGTTGG	1860
Qy	1861	GGCACTGCAAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCCTGCCAAGATGCTCAA	1920
Db	1861	GGCACTGCAAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCCTGCCAAGATGCTCAA	1920
Qy	1921	CTGAGCCCAACTTGTGTGATGATGAGGAGGATTTGAGAGCGTCTTCGGTGGGGAACCG	1980
Db	1921	CTGAGCCCAACTTGTGTGATGATGAGGAGGATTTGAGAGCGTCTTCGGTGGGGAACCG	1980
Qy	1981	TTGAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019	
Db	1981	TTGAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019	

RESULT 2	
AX137736	
LOCUS	2019 bp
DEFINITION	Sequence 1 from Patent EP1077260.
ACCESSION	AX137736
VERSION	AX137736.1 GI:14273909
	linear PAT 30-MAY-2001

KEYWORDS	Mice minute virus	Db	481	TGTAATGTGCAACTAACACGAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAAGCAAT	540
SOURCE	Mice minute virus	Qy	541	GAGTGGTTACTTACTTACTTACTTATATAGCATATAGCAACCAAAAGACATATACCAAGTGT	600
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	Db	541	GAGTGGTTACTTACTTACTTACTTATATAGCATATAGCAACCAAAAGACATATACCAAGTGT	600
REFERENCE	1 Nueesch, J. and Rommelaere, J.	Qy	601	GTTCCTTTTGGAAACATGATGCTTACTTACTTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
AUTHORS	Parvovirus nsl variants	Db	601	GTTCCTTTTGGAAACATGATGCTTACTTACTTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
TITLE	Patent: EP 1077260-A 1 21-FEB-2001;	Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)	Db	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
FEATURES	Location/Qualifiers	Qy	721	AAAGAAGGAGCGGCCCATCTAGTGAGCAAACTATATACATGATGACATGGGCCAGAAACG	780
source	1. .2019	Db	721	AAAGAAGGAGCGGCCCATCTAGTGAGCAAACTATATACATGATGACATGGGCCAGAAACG	780
	/organism="Mice minute virus"	Qy	781	GTTGAAACACACAGTAAACCACTGCGCAGGAAACTTAAGCGGCGAGAAATTCAAACTAAAAAA	840
	/mol_type="unassigned DNA"	Db	781	GTTGAAACACACAGTAAACCACTGCGCAGGAAACTTAAGCGGCGAGAAATTCAAACTAAAAAA	840
	/db_xref="taxon:10794"	Qy	841	GAAAGTTTCTATTAAAACTACACTTAAGAGCTGGTGCAATAAAGAGTAACTCACCAGAG	900
	/notes="Wildtype Parvovirus NS1"	Db	841	GAAAGTTTCTATTAAAACTACACTTAAGAGCTGGTGCAATAAAGAGTAACTCACCAGAG	900
	1. .2019	Qy	901	GACTGATGATGATGAGCAGACAGTACATTGAAATGATGCTCAACAGGTGGAGAA	960
	/notes="unnamed protein product"	Db	901	GACTGATGATGATGAGCAGACAGTACATTGAAATGATGCTCAACAGGTGGAGAA	960
	/codon_start=1	Qy	961	AACTCTCTGAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA	1020
	/protein_id="CAC39989.1"	Db	961	AACTCTCTGAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA	1020
	/db_xref="GI:14273910"	Qy	1021	TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
	/translations="MAGNAYSDVLGATNWLKEKSQEVFVFKNENVQLNGKQIDGM NSYKLEQDELQSGAETWQSEDMEWETTVDEMTKQVFI FDSLVKKCLFVL NTKNIPEGDVNWFVQHEWKGQDGHCHVLIGKDFSOAGCKWRRLNLYVWSRLVTA ARTKADFLLEKAETS KLTNFSLPDTRTCRI FAFHWNVYKVCHKAI CCVLNRQGGKR NTVLFGPASTGKSI IAOIAOAVGNVCYNAANVFPFNDCTNKL I WVEEAGNFGQ OVNOFKAI CSGOTIRIDOKGKSKQI EPTFVIMTNTENITVVI GCEERPEHTQPIRD RMLNIHLTHLPGDFGLVDKNWPMI CAMLVKNGYQSTWAS YCARKNGKVPDWSWAE KVPYTPINLIGLSRFFTPKTSQNSYALTPLEADLELALPEWSTNPTNPVAGTAE TQNTGAGSKACQDQLSPTWSBIEBDRACFGAEPLKKDFSEPLND"	Db	1021	TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
ORIGIN		Qy	1081	ACAAGAACCTGCAAGATTTTGTGCTTTCATGCTGGCACTATGTTAAAGTTGCGCATGCT	1140
	Query Match 99.98; Score 2017.4; DB 6; Length 2019;	Db	1081	ACAAGAACCTGCAAGATTTTGTGCTTTCATGCTGGCACTATGTTAAAGTTGCGCATGCT	1140
	Best Local Similarity 100.0%; Pred. No. 0;	Qy	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTTCATGGACCA	1200
	Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Db	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTTCATGGACCA	1200
		Qy	1201	GCAGACACAGGCAAACTCTATTATTGCACAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
		Db	1201	GCAGACACAGGCAAACTCTATTATTGCACAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
		Qy	1261	TGCTATAATGCAAGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTGATT	1320
		Db	1261	TGCTATAATGCAAGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTGATT	1320
		Qy	1321	TGGGTAGAGAGAGCTGGTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380
		Db	1321	TGGGTAGAGAGAGCTGGTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380
		Qy	1381	GGTCAAGCTATTTCGCATTGATCAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA	1440
		Db	1381	GGTCAAGCTATTTCGCATTGATCAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA	1440
		Qy	1441	GTCAATCATGACCAAAATGAGAACTTATACAGTGGTTCAGAAATAGGCTGCGAAGAAAGCA	1500
		Db	1441	GTCAATCATGACCAAAATGAGAACTTATACAGTGGTTCAGAAATAGGCTGCGAAGAAAGCA	1500
		Qy	1501	GACACACTCAACCAATCGAGACAGAAATGCTTAACTTCAATCTTAACATACCTTGCCT	1560
		Db	1501	GACACACTCAACCAATCGAGACAGAAATGCTTAACTTCAATCTTAACATACCTTGCCT	1560
		Qy	1561	GGTGAATTTGGTTGGTGGCAAAAATGATGCCCCCATGATTGCTGCTGGTAAAG	1620
		Db	1561	GGTGAATTTGGTTGGTGGCAAAAATGATGCCCCCATGATTGCTGCTGGTAAAG	1620



561 GTGCTTAAACAAGAAATATATTTCTGGTGTATGTTAAATGGTTTGTGCAACATGATGG 620  
361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGAACTTTAGTCAAGCT 420  
621 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGAACTTTAGTCAAGCT 680  
421 CAAGGGAATGGTGGAGAGGCACTAATATTTTACTGGAGCAGATGGTTGTTAAAGCC 480  
681 CAAGGGAATGGTGGAGAGGCACTAATATTTTACTGGAGCAGATGGTTGTTAAAGCC 740  
481 TGTAAATGTGCAACTAACAACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAGACAAT 540  
741 TGTAAATGTGCAACTAACAACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAGACAAT 800  
541 GAGTGGGTTACTCTACTTACTTATAGCAATAAGCAAAACCAAAAGCACTATACCAAGTGT 600  
801 GAGTGGGTTACTCTACTTACTTATAGCAATAAGCAAAACCAAAAGCACTATACCAAGTGT 860  
601 GTTCTTTTGGAAACATGATGCTTACTTACTTATTTTAACTAAAGAAAGAAATAAGCACTAGT 660  
861 GTTCTTTTGGAAACATGATGCTTACTTACTTATTTTAACTAAAGAAAGAAATAAGCACTAGT 920  
661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAACTTAACTTTTAA 720  
921 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAACTTAACTTTTAA 780  
721 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 780  
981 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 1040  
781 GTTGAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAACTAATAAATAA 840  
1041 GTTGAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAACTAATAAATAA 1100  
841 GAAGTTCTTATTAATACTACATTAAGAGCTGGTGCATTAAGAGAGTAACTCACCAGAG 900  
1101 GAAGTTCTTATTAATACTACATTAAGAGCTGGTGCATTAAGAGAGTAACTCACCAGAG 1160  
901 GACTGGATGATGATGACGCGCAGACAGTTACATTTGAATGATGCTCAACAGGCTGGAGAA 960  
1161 GACTGGATGATGATGACGCGCAGACAGTTACATTTGAATGATGCTCAACAGGCTGGAGAA 1220  
961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
1221 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1280  
1021 TTTGACTTAATTTAGAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCTGAC 1080  
1281 TTTGACTTAATTTAGAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCTGAC 1340  
1081 ACAAGAACCTGCAAGATTTTGTCTTTTCTGCTGGAATCTATGTTAAAGTTTGCCATGCT 1140  
1341 ACAAGAACCTGCAAGATTTTGTCTTTTCTGCTGGAATCTATGTTAAAGTTTGCCATGCT 1400  
1141 ATTTGCTGTTTTTAAACAGCAAGGAGGCAAGAAATACTGTTTATTTATGAGCA 1200  
1401 ATTTGCTGTTTTTAAACAGCAAGGAGGCAAGAAATACTGTTTATTTATGAGCA 1460  
1201 GCAGCAGCAGGCAATCTATTTATGCAAGCCATAGCAACAGCTTGGCAATGTTGTT 1260  
1461 GCAGCAGCAGGCAATCTATTTATGCAAGCCATAGCAACAGCTTGGCAATGTTGTT 1520  
1261 TGCTATAATCAGCCCAATGTAATTTTCCATTTTAAAGTGTACCAACAGAACTTGATT 1320  
1521 TGCTATAATCAGCCCAATGTAATTTTCCATTTTAAAGTGTACCAACAGAACTTGATT 1580  
1321 TGGGTAGAAGAGCTGGTAACTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380  
1581 TGGGTAGAAGAGCTGGTAACTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1640  
1381 GGTCAAGCTATTCGATTTGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440

1641 GGTCAAACTATTGCGATTGATCAAAAGGAAAGGAGCAAGCAAGATTGAACCAACACCA 1700  
1441 GTCATCATGACCAACAATGAGAACATTACAGTGGTTCAGATAGGCTGCGAAGAAACCA 1500  
1701 GTCATCATGACCAACAATGAGAACATTACAGTGGTTCAGATAGGCTGCGAAGAAACCA 1760  
1501 GAAACACATCAACCAATGAGAGACAGATGCTTAACTTATCATCTAACACATCTTGGCT 1560  
1761 GAAACACATCAACCAATGAGAGACAGATGCTTAACTTATCATCTAACACATCTTGGCT 1820  
1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTGCTTGGTTGGTAAAG 1620  
1821 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTGCTTGGTTGGTAAAG 1880  
1621 AATGTTTCAATCTACCATGAGCAAGCTACTGCTTAAATGGGCAAGTTCCTGATTGG 1680  
1881 AATGTTTCAATCTACCATGAGCAAGCTACTGCTTAAATGGGCAAGTTCCTGATTGG 1940  
1681 TCAGAAAACTGGCGGAGCAAGGTCGCAACTCTCTATATAATTTACTAGGTTGCGACGC 1740  
1941 TCAGAAAACTGGCGGAGCAAGGTCGCAACTCTCTATATAATTTACTAGGTTGCGACGC 2000  
1741 TCACCAATTCACGACACCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTTAATCCACTT 1800  
2001 TCACCAATTCACGACACCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTTAATCCACTT 2060  
1801 GATCGGATCTCGAGGACCTGCTTTAGAGCCTTTGAGGACACACCAATACTCTCTGTTGG 1860  
2061 GATCGGATCTCGAGGACCTGCTTTAGAGCCTTTGAGGACACACCAATACTCTCTGTTGG 2120  
1861 GGCACCTGCAAGAACCCAGAACACTGCGGAGCTGGTTCCAAAGCCGCAAGATGCTCA 1920  
2121 GGCACCTGCAAGAACCCAGAACACTGCGGAGCTGGTTCCAAAGCCGCAAGATGCTCA 2180  
1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACCG 1980  
2181 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACCG 2240  
1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019  
2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2279

## RESULT 4

LOCUS MVMPGCG 5149 bp ss-DNA linear VRL 22-MAY-1995  
DEFINITION Minute virus of mice, complete genome.  
ACCESSION J02275.1 M12521 M14704  
VERSION J02275.1 GI:332293  
KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.  
SOURCE Mice minute virus  
ORGANISM Mice minute virus  
REFERENCE 1 (bases 1 to 5149)  
AUTHORS Astell,C.R., Thomson,M., Merchinsky,M. and Ward,D.C.  
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus  
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)  
MEDLINE 83143341  
PUBMED 6298737  
REFERENCE 2 (bases 1 to 5149)  
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.  
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain J. Virol. 57 (2), 656-669 (1986)  
JOURNAL MEDLINE 86115415  
PUBMED 3502703  
REFERENCE 3 (sites)  
AUTHORS Morgan,W.R. and Ward,D.C.  
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs

J. Virol. 60 (3), 1170-1174 (1986) 87061199 3783817 PUBMED	J. Virol. 60 (3), 1170-1174 (1986) 87061199 3783817 PUBMED	Original source text: Minute virus of mice (strain MVM(p)), passed in mouse 1 (variant A-9) cells. The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group. The sequence below corresponds to the plus (+) strand, also referred to as the C-strand. The minus (-) strand is also referred to as the V-strand. The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1]. The left and right halves of the genome encode two distinct, but overlapping transcriptional units. The transcripts can be summarized [1] (1 map unit (mu) = 51 bp): R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu R3 is the major transcript. There are two major open reading frames, both on the plus (or C) strand. The left side ORF (261-2279) probably encodes a non-capsid protein of 85 kd; the right side ORF probably encodes the viral capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known. revision 4804 4870 a-65bp-a in [2]; aa in [1] [2] sites [1]. [3] sites; splice sites. Location/Qualifiers 1. 5149 /organism="Mice minute virus" /mol_type="genomic DNA" /strain="MVM(p)" /db_xref="taxon:10794" /lab_host="mouse 1 (variant A-9) cell" 114. 2279 /gene="NS1" 114. 2279 /gene="NS1" /notes="putative" /codon_start=1 /product="nonstructural protein" /protein_id="AAA67108.1" /db_xref="GI:825481" /translations="MISGSGSLNQAKRWAFKVKYKOLLKSVYLFHFSVSRDAOKE SNOLTMAGNAYSDEVIGATNWLKEKSNQEVFVFKNENVLQNGKDGHNYSKKELQ DELKSLQGAETDQSEMEWETTVDGMTKQVFIQDLSLVKKCLFVNTKNIFFPD VNMVQHEGKQDQCHVLIQGGKFSQAQKQWRRQLNYSRWLVLTACNVLTPAE RIKLRIAEENWVLLTYKQKDYTKCVLFQNMIAFYFLTKKISIPSPDGGY FLSSDGKNTFLKEGHEHLVSKLYTDMRPETVTTVTAQETKGRITQKKEVSIK TTLKELVHKVTSPEDDMMQPDYSYIEMMAQPGENLLKNTLICTLTARTAEFL ILKATSKLNTSLPDTCTCRIFAFHGWNVKVCVHAICCVLRQGGKNTLVFHGPA STGKSIQAIAQAVGNCYNAANVFPNDCTNKLWVEEAGNFQGVQVQFKAIC SGQTIIRDQKGSQKIEPTVIMTNENITVVRIGCEERPEHTQPIRDMNLHLTH TLPGDFGLVDKNEWPMICAWLVKNGYSTWASYCAWKGVKPDWSENWAPKVPPTNL LGSARSPFTTPKSTPLSONYALTPLASDLEDALEPWTPTVAGTAETQNTGAGS KACQDQLSTWSEIREDLRACFGAEPLKKDFSEPLNLD" 200. 2279 /gene="NS1" 261. 2279 /gene="NS1" /codon_start=1 /product="nonstructural protein" /protein_id="AAA67109.1" /db_xref="GI:332295" /translation="MAGNAYSDEVLGATNWLKEKSNQEVFVFKNENVLQNGKDIGW	NSYKKELODELKSLQGAETTWQSEDMEWETTVDGMTKQVFIQDLSLVKKCLFVIL NTKNIFPDGVNFWHEGKQDQCHVLIQGGKFSQAQKQWRRQLNYSRWLVTA CNVQLTAPARIKLRIAEENWVLLTYKQKDYTKCVLFQNMIAFYFLTKKIS TPSPDGGYFLSSDGKNTFLKEGHEHLVSKLYTDMRPETVTTVTAQETKGRIR QTKKEVSIKNTLKLKELVHKVTSPEDDMMQPDYSYIEMMAQPGENLLKNTLICTLT ATKTAPDLILSKAETSKLNTSLPDTCTCRIFAFHGWNVKVCVHAICCVLRQGGKR NTLVFHGPASTGKSIQAIAQAVGNCYNAANVFPNDCTNKLWVEEAGNFQGV QNVQKAIQSGQTIIRDQKGSQKIEPTVIMTNENITVVRIGCEERPEHTQPIR RMLNLHLTHLPGDFGLVDKNEWPMICAWLVKNGYSTWASYCAWKGVKPDWSENWAE PKVPPTNLGSGARSPFTTPKSTPLSONYALTPLASDLEDALEPWTPTVAGTAET TQNTGAGSKACQDQLSTWSEIREDLRACFGAEPLKKDFSEPLNLD" 2002. 2398 /gene="VP" 2002. 2280 /gene="VP" /note="major transcription start site" /number=1 2006. 2280 /gene="VP" /note="minor transcription start site" /number=1 2009. 2280 /gene="VP" /notes="minor transcription start site" /number=1 2281. 2398 /gene="VP" /note="alternative intron" 2281. 2376 /gene="VP" /note="alternative intron" join(2286. 2316,2399. 4557) /gene="VP1" join(2286. 2316,2399. 4557) /gene="VP1" /codon_start=1 /protein_id="AAA67111.1" /db_xref="GI:332294" /translation="MAPPAKRAKRWVPPGYKYLPGNSLDQGEPTNPDAAKEHDE AYDOVIKSGKNPLYFSAADQRFIDQTKADQWGGKVGHYFRTKRAFAPKLATSE GTSVSRAGKTRPPAYIFINQARAKKLTSSAAQSSQTSMDGTSPDGNVAHSA RVERAADGGSGGGGGGSGVSTGSDNQTHTFRFLDGHVEITALATRLVHLNMP KSENYCRIRVHNTDITSVKNWAKDAHEQIWPWSLVNDAWGWLQSDWQYICNT MSQLNLVSLDQEIFNVVLKTVTEQDLGGQAIKIYNNDLTACMVAVDSNNILPYTPAA NSMETIGFYPMKPTIASPYRYFCVDRDLSVTYENQEGTVENHVMGTPKGMSQFTI ENTQOITLLRTGDEFATGYTFDTNSVKLTHTVQTNROLQGPPLSTFPEATDAGTL TAQGSRHGTTQMVANVSEAIRPAQVGCQPHNDFEASRAGPFAAPKVPADITQGV DKRANGSVRYSGKHGHNWASHGAPERYTWDETSGSGRDTKGFQISAPLVPPP LNGILTNANPIGTNDIHFSNVFNSYGLPTAFSPSPVYPOGQIMDKDLBHKPLRH ITAPFYCKNAFQOMLVRLGPLNLTQYDPNGATLSRIVTYGTFWKGKLTMRKLRAN TTWNPVYQVSAEDNGNSYMSVTKLPTATGNNQSVPLITRPVARTY" 2286. 2354 /gene="VP3" /note="ORF1; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67110.1" /db_xref="GI:825482" /translation="MAPPAKRAKRGKGLRDGWLGVY" 2286. 2316 /gene="VP" /number=1 2317. 2398 /gene="VP" /notes="VP intron (alt.)" 2332. 2361 /gene="VP" /note="ORF3; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67112.1" /db_xref="GI:825483" /translation="MVGWVGINV"
---	---	---	---

CDS	2354..2398 /gene="vp" /note="ORF2; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67113.1" /db_xref="GI:825484" /translation="MFNYLFYRPBITWF" 2399..>4557 /gene="vpi" /number=2 2794..4557 /gene="vpi1" /note="vp2" /codon_start=1 /protein_id="AAA67114.1" /db_xref="GI:33296" /translation="MSDGTSPQSDGNVHSAARVERAADPGSGGGSGGGGCVGVST GSYNDQITPLRGDWVEIETALATRLVHLNMPSENCRIIRVHTNTDTISVKGNMAKOD AHEQIWTPWSLWDANAGWVLQSDQYICNTMSQLNLVDLOEIFNVILKTVTQEAD GGQAIKTYENDLTACMMVAVDNSNNILPYTPAANSMETLGYPWKPTTASPRYFYCFVD RDLSVTYENOEGTVHNVMGTPKGMNQFFTIENTQOITLLRTGDDEATGYVPDINS VKLTHTWQNRQIGPOLLSFTPEADTDAGTLTAQGRHGTTQMVGUNVSEAIRTRPA QVGEQHNFDFESRAGPAPKPVADIOTGVDEKANGSVRYSKOKHGENWASHGPAA PERYTQDETSGSRDTKDGFIQSAPLIIVPPPLNGILTNPAGTKNDIHFSNVFNYS GPLTAFSHPSVPVPOGGIWMKELDEHKRLHITAPEVCVRKNAPQQMLVLRGLNPLTDQ
exon	
CDS	
Query Match	99.9%; Score 2017.4; DB 14; Length 5149;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2018;	Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 ATGCCTCGGAATGCTTACTCTCATGATGAAGTTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db	261 ATGCCTCGGAATGCTTACTCTGATGAGTTTTTGGGAGCAACCAACTGGTTAAAGGAAAA 320
Qy	61 AGTAAACAGGAAGTGTTCTCATTTGTTTTTAAAAATGAAAATGTTTCAACTGAATGGAAAA 120
Db	321 AGTAAACAGGAAGTGTTCTCATTTGTTTTTAAAAATGAAAATGTTTCAACTGAATGGAAAA 380
Qy	121 GATATCCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA 180
Db	381 GATATCCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA 440
Qy	181 CGAGGAGCGGAAACTACTTTGGGACCAGAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 240
Db	441 CGAGGAGCGGAAACTACTTTGGGACCAGAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 500
Qy	241 GAAATGACCAAAAAAGCAAGTATTCATTTTGAATTCTTTGGTTAAAAAATGTTTATTGTA 300
Db	501 GAAATGACCAAAAAAGCAAGTATTCATTTTGAATTCTTTGGTTAAAAAATGTTTATTGTA 560
Qy	301 GTGCTTTAACACAAAGATAATTTTCCGTGCATGTTTAATTCGTTTGTGCAACATGGAATGG 360
Db	561 GTGCTTTAACACAAAGATAATTTTCCGTGCATGTTTAATTCGTTTGTGCAACATGGAATGG 620
Qy	361 GGAAAAAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT 420
Db	621 GGAAAAAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT 680
Qy	421 CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC 480
Db	681 CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC 740
Qy	481 TGTAAATGTCAACTTAACACAGCTGTAAGAAATTAACCTTAACAGAAATAGCAGAGACACAAT 540
Db	741 TGTAAATGTCAACTTAACACAGCTGTAAGAAATTAACCTTAACAGAAATAGCAGAGACACAAT 800
Qy	541 GAGTGGGTTACTCTACTTACTTATAAGCATAGCAAAACCAAAAAAGACTATACCAAGTGT 600
Db	801 GAGTGGGTTACTCTACTTACTTATAAGCATAGCAAAACCAAAAAAGACTATACCAAGTGT 860
Qy	601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTTTAACTAAAAAGAAATTAAGCACTAGT 660

Db	861	GTTC	TTTTTTGG	AAACATG	ATGCTT	ACTATTTT	TTTAACT	AAAAAG	AAAAAT	AAGCACT	AGT	920					
Qy	661	CCAC	CAAGAG	CAGG	AGGCT	ATTTTCT	TAGCAGT	GACTCT	GGCTG	GAAAACT	TAACCTTTT	720					
Db	921	CCAC	CAAGAG	CAGG	AGGCT	ATTTTCT	TAGCAGT	GACTCT	GGCTG	GAAAACT	TAACCTTTT	980					
Qy	721	AAAG	AGGCGAG	CGCC	ATCT	AGTGAG	CAAACT	ATAC	ACTG	ATG	CATCGGCGCA	AGAAACG	780				
Db	981	AAAG	AGGCGAG	CGCC	ATCT	AGTGAG	CAAACT	ATAC	ACTG	ATG	CATCGGCGCA	AGAAACG	1040				
Qy	781	GTG	GAAC	CACAC	GTAACT	CGCGC	AGAACT	TAAGCG	CGGAG	AACTTCAA	AACTTAAAAA	840					
Db	1041	GTG	GAAC	CACAC	GTAACT	CGCGC	AGAACT	TAAGCG	CGGAG	AACTTCAA	AACTTAAAAA	1100					
Qy	841	GAAG	TTTTCT	ATAAAAA	CTAC	ACTTAA	AGAGCT	GGTG	CAATA	AAAAAG	ATAACCTC	ACAGAG	900				
Db	1101	GAAG	TTTTCT	ATAAAAA	CTAC	ACTTAA	AGAGCT	GGTG	CAATA	AAAAAG	ATAACCTC	ACAGAG	1160				
Qy	901	GACT	GTG	ATG	ATG	CAGC	CAGCTT	CATTTG	AAATG	ATG	CTCAAC	AGGTGG	960				
Db	1161	GACT	GTG	ATG	ATG	CAGC	CAGCTT	CATTTG	AAATG	ATG	CTCAAC	AGGTGG	1220				
Qy	961	AAC	CTGCTG	AAAAAT	ATCG	CTAG	AGATTT	GTG	TAC	ACTTCT	TAG	CCAGAA	CCAAAAC	1020			
Db	1221	AAC	CTGCTG	AAAAAT	ATCG	CTAG	AGATTT	GTG	TAC	ACTTCT	TAG	CCAGAA	CCAAAAC	1280			
Qy	1021	TTTG	ACTTAA	TTTTT	TAGAAAA	AGCTG	AAAC	CAGCA	AACTAA	ACCACTTTT	TCTC	CTCCTG	AC	1080			
Db	1281	TTTG	ACTTAA	TTTTT	TAGAAAA	AGCTG	AAAC	CAGCA	AACTAA	ACCACTTTT	TCTC	CTCCTG	AC	1340			
Qy	1081	ACA	AGAAC	CTG	CAGAA	TTTTT	TGCTTTT	CATG	GGCTG	GAACTAT	GT	TAAAGTTT	TG	1140			
Db	1341	ACA	AGAAC	CTG	CAGAA	TTTTT	TGCTTTT	CATG	GGCTG	GAACTAT	GT	TAAAGTTT	TG	1400			
Qy	1141	ATT	TGCTG	TGCTTTT	TAAAC	CACAG	CAGG	CAAAAG	AAAT	ACTGTTT	TATTTT	CATGG	CA	1200			
Db	1401	ATT	TGCTG	TGCTTTT	TAAAC	CACAG	CAGG	CAAAAG	AAAT	ACTGTTT	TATTTT	CATGG	CA	1460			
Qy	1201	GCC	AGCA	CAGG	CAAACT	CTATTT	TGCA	CAAG	CCATAG	CACAA	G	CAGTTG	GCAATG	1260			
Db	1461	GCC	AGCA	CAGG	CAAACT	CTATTT	TGCA	CAAG	CCATAG	CACAA	G	CAGTTG	GCAATG	1520			
Qy	1261	TGCT	ATA	TG	CAG	CCAA	TGTA	AACTTTT	TCCATTT	TAATG	ACTGT	TACCA	AGAACTT	1320			
Db	1521	TGCT	ATA	TG	CAG	CCAA	TGTA	AACTTTT	TCCATTT	TAATG	ACTGT	TACCA	AGAACTT	1580			
Qy	1321	TGG	GTAGA	CAAG	CTGT	TAACTTT	TG	GACAG	CAAGTAA	CCAGTTT	TAAAG	CCATTT	TGCTCT	1380			
Db	1581	TGG	GTAGA	CAAG	CTGT	TAACTTT	TG	GACAG	CAAGTAA	CCAGTTT	TAAAG	CCATTT	TGCTCT	1640			
Qy	1381	GGT	CAAG	CTATTT	CGC	ATTG	TCAAAAA	AGGAAAA	GGCAG	CAAA	CAGATTG	GAACCA	ACCA	1440			
Db	1641	GGT	CAAG	CTATTT	CGC	ATTG	TCAAAAA	AGGAAAA	GGCAG	CAAA	CAGATTG	GAACCA	ACCA	1700			
Qy	1441	GTCA	T	ATG	ACCA	CAAA	TG	AGAC	ATTTAC	AGTGT	CGAAT	TAG	GCTCGG	AGAAAG	1500		
Db	1701	GTCA	T	ATG	ACCA	CAAA	TG	AGAC	ATTTAC	AGTGT	CGAAT	TAG	GCTCGG	AGAAAG	1760		
Qy	1501	GA	CAC	ACTC	CA	CCAA	T	CG	AG	CAG	AAATG	CTTAA	CA	CATAC	CTT	1560	
Db	1761	GA	CAC	ACTC	CA	CCAA	T	CG	AG	CAG	AAATG	CTTAA	CA	CATAC	CTT	1820	
Qy	1561	GGT	G	ACTTTT	TG	TTT	TG	CAAAAA	T	GAAT	TGG	CCCAT	TG	TTT	TG	1620	
Db	1821	GGT	G	ACTTTT	TG	TTT	TG	CAAAAA	T	GAAT	TGG	CCCAT	TG	TTT	TG	1880	
Qy	1621	AAT	GGTTA	CA	AA	TCTAC	CAT	GG	CAAG	CTAC	TG	TG	CT	AAAT	TGG	1680	
Db	1881	AAT	GGTTA	CA	AA	TCTAC	CAT	GG	CAAG	CTAC	TG	TG	CT	AAAT	TGG	1940	
Qy	1681	TC	AGAA	AACT	TGG	CGG	AGCC	AA	AGG	TG	CC	AACT	TAATTTT	ACTAG	GTTCG	CAAGC	1740
Db	1941	TC	AGAA	AACT	TGG	CGG	AGCC	AA	AGG	TG	CC	AACT	TAATTTT	ACTAG	GTTCG	CAAGC	2000



QY 1741 TCACCAATTCAGCACCGAAAGTACCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 2001 TCACCAATTCAGCACCGAAAGTACCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060  
QY 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAAAATATCTCTGTTGG 1860  
Db 2061 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAAAATATCTCTGTTGG 2120  
QY 1861 GGCACTCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCTGSCCAAGATGGTCAA 1920  
Db 2121 GGCACTCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCTGSCCAAGATGGTCAA 2180  
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGTTGGGGAACCG 1980  
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGTTGGGGAACCG 2240  
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2019  
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2279

RESULT 5  
AX137739  
LOCUS  
DEFINITION  
Sequence 4 from Patent EPI077260.  
ACCESSION  
AX137739  
VERSION  
AX137739.1 GI:14273913  
KEYWORDS  
Mice minute virus  
SOURCE  
Mice minute virus  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
REFERENCE  
1  
Nuesch, J. and Rommelaere, J.  
Parvovirus ns1 variants  
Patent: EP 1077260-A 4 21-FEB-2001;  
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts  
(DE)

FEATURES  
source  
1..2019  
/organism="Mice minute virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10794"  
/notes="parvovirus ns1 variant"  
1..2019  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC39991.1"  
/db\_xref="GI:14273914"

CDS  
1..2019  
/translations="MAGNAYSDEVLGATNWLKESNQEVFSFVFKNENVQLNGKDIGW  
NSYKLEODELSLORGAEFTWDQEDMEWETTVDEMFKKOFIFDSLKVKCLPEVL  
NTKNIIPGDVNVFQHWGDKQCHVLIGKDFSOAGKWREROLNVYWSRWLTA  
CNVLPFAERIKUREIAEDNEWTLITYRKHQKDYTKCVLPGNMIAYFLTKKIS  
TSPRRDGGVPLSDSGWKNFLEKGERHLVSKLYTDDMRPEVETVTTAQTETKGR1  
QTKVEAIKTKLEKLVKRVTSPEWMMQPDYSYIEMMAQPGGENLLKNTLETCTLT  
ARTKAFDLILEKAETSLSLTPDTRICRIFAFGWNVKVKHAIICVLRQGGKR  
NTVLFGPASTGKSIITAOAIVAGNVGCVYNAANVFPENDCTNKLIVVEBAGNFGQ  
QVNFKNICSGQIRIDQKSGSKQIEPTVIMTNENITVVRIGCEEPETHQPLRD  
RMLNHLTHLPDGLVDLKNWPMICAMLVKNGYSTWASYCAKMGKVPDWSNWA  
PKVPTPINLIGSARSPTPKPSQNVALTPLASDLDELALPEWSTNTPVAGTAE  
TONTGAGSKACQDGLSPTWSIEIEDLRACFGAEPLKKDFSEPLND"

ORIGIN  
Query Match 99.88; Score 2015.8; DB 6; Length 2019;  
Best Local Similarity 99.98; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 1 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
QY 61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120

Db 61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCGAGCTGAAATCTTTTACAA 180  
Db 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCGAGCTGAAATCTTTTACAA 180  
QY 181 CGAGGAGCGGAAAACTACTCTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240  
Db 181 CGAGGAGCGGAAAACTACTCTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240  
QY 241 GAATACCAAAAAAGCAAGTATTCATTTTGGTTTGAATCTTCTTGGTTAAAAAATGTTTATTGAA 300  
Db 241 GAATACCAAAAAAGCAAGTATTCATTTTGGTTTGAATCTTCTTGGTTAAAAAATGTTTATTGAA 300  
QY 301 GTGCTTTAACACAAAGAAATATATTTCTCGTGTGATGTTAAATTTGGTTTGTGCAACATGAATGG 360  
Db 301 GTGCTTTAACACAAAGAAATATATTTCTCGTGTGATGTTAAATTTGGTTTGTGCAACATGAATGG 360  
QY 361 GGAAAAAGCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTAGTCAAGCT 420  
Db 361 GGAAAAAGCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTAGTCAAGCT 420  
QY 421 CAAGGAAATGTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGTTAAACAGCC 480  
Db 421 CAAGGAAATGTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGTTAAACAGCC 480  
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGCAAT 540  
Db 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 540  
QY 541 GAGTGGTTTACTCTACTTATAGCAATAAGCAAAACCAAAAGACTATATACCAAGTGT 600  
Db 541 GAGTGGTTTACTCTACTTATAGCAATAAGCAAAACCAAAAGACTATATACCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATATAGCACTAGT 660  
Db 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATATAGCACTAGT 660  
QY 661 CCACCAAGAGACGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720  
Db 661 CCACCAAGAGACGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720  
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 780  
Db 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 780  
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTAAGCGCGCAGANTTCAAACTAAAAA 840  
Db 781 GTTGAACACACAGTAACCACTGCGCAGGAAACTAAGCGCGCAGANTTCAAACTAAAAA 840  
QY 841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGATATAAAAGAGTAACCTCACCAGAG 900  
Db 841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGATATAAAAGAGTAACCTCACCAGAG 900  
QY 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAAATGATGGCTCAACACAGTGGAGAA 960  
Db 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAAATGATGGCTCAACACAGTGGAGAA 960  
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020  
Db 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020  
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACAGCAAACTAACTAACTTTTCACTGCTCTGAC 1080  
Db 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACAGCAAACTAACTAACTTTTCACTGCTCTGAC 1080  
QY 1081 ACAAGAACCTGAGAAATTTTCTTTTCATGCTGGAACTATGTTTAAAGTTTGCCTAGCT 1140  
Db 1081 ACAAGAACCTGAGAAATTTTCTTTTCATGCTGGAACTATGTTTAAAGTTTGCCTAGCT 1140  
QY 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTATTCATGACCA 1200  
Db 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTATTCATGACCA 1200



Qy	1201	GCAGCAGCGAAATCTATTATTGCAAGGCCATAGCAAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCAGCAGCGAAATCTATTATTGCAAGGCCATAGCAAGCAGTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATCAGCGCAATGTAATCTTTCCATTTAATGACGTGTACCAACAGAACTTGATT	1320
Db	1261	TGCTATAATCAGCGCAATGTAATCTTTCCATTTAATGACGTGTACCAACAGAACTTGATT	1320
Qy	1321	TGGGTAGAAGAGCTGGTAATCTTTGGACAGCAAGTAAACAGATTTAAAGCAATTTGCTCT	1380
Db	1321	TGGGTAGAAGAGCTGGTAATCTTTGGACAGCAAGTAAACAGATTTAAAGCAATTTGCTCT	1380
Qy	1381	GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGAGCGAGCAAAACAGATTGAACCAACCA	1440
Db	1381	GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGAGCGAGCAAAACAGATTGAACCAACCA	1440
Qy	1441	GTCTATCATGACCAAAATGAGAACATTTACAGTGTGTGAGATAGCTGCGAAGAAAGACCA	1500
Db	1441	GTCTATCATGACCAAAATGAGAACATTTACAGTGTGTGAGATAGCTGCGAAGAAAGACCA	1500
Qy	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTTCACTTAACACATACCTTGCT	1560
Db	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTTCACTTAACACATACCTTGCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTGCTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTGCTTGGTAAAG	1620
Qy	1621	AATGGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG	1680
Db	1621	AATGGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG	1680
Qy	1681	TCAGAAAACCTGGGCGAGCCAAAGGTGCTTAATTTACTAGTTTGGCAGCG	1740
Db	1681	TCAGAAAACCTGGGCGAGCCAAAGGTGCTTAATTTACTAGTTTGGCAGCG	1740
Qy	1741	TCACCATTTACGACACCGAAAAGTACGCTCTCAGCGAGAACTATGCACTAACCTCCACTT	1800
Db	1741	TCACCATTTACGACACCGAAAAGTACGCTCTCAGCGAGAACTATGCACTAACCTCCACTT	1800
Qy	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATCTCTCTGTTGG	1860
Db	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATCTCTCTGTTGG	1860
Qy	1861	GGCACTGCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Db	1861	GGCACTGCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Qy	1921	CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGCGGTGCTTCGGTGGGAACCG	1980
Db	1921	CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGCGGTGCTTCGGTGGGAACCG	1980
Qy	1981	TTCAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019
Db	1981	TTCAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019

RESULT 6	
AX137743	
LOCUS	AX137743
DEFINITION	Sequence 8 from Patent EP1077260.
ACCESSION	AX137743
VERSION	AX137743.1
KEYWORDS	GI:14273917
SOURCE	Mice minute virus
ORGANISM	Mice minute virus
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
AUTHORS	1
TITLES	Nuesch, J. and Rommelaere, J.
JOURNAL	Parvovirus ns1 variants
	Patent: EP 1077260-A 8 21-FEB-2001;
	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

FEATURES	(DB)
source	Location/Qualifiers
	1. .2019
	/organism="Mice minute virus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:10794"
	/notes="Parvovirus ns1 variant"
CDS	1. .2019
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC39993.1"
	/db_xref="GI:14273918"
	/translation="MAGNAYSDEVLTGATNWLKEKSQEVFSPVKENVQVNGKDIGH
	NSYKELQDELKSLQRGAEITTDQSEMEWETTVDEMTKKQVFIIDSLVKKCLFVL
	NTKNTIPGVNMFVQHWGKQDQGHVLIIGKDFQAOQKWRRLNLYWSRLVTA
	CNVQLTPABRIKLREIAEDNEWTLTYKHQTKDYTKCVLFKNMIAIYVFLTKKRTI
	TSPRDGGVPLSDSKWNTFLKEGRLHLSVSKYTDMDRPETVETVTTAQETKRGRI
	QTKKEYSIKTTLKELVHKVTSPEMMQPDYSIEMMAQPGGGENILLKNTLEICTLTL
	ARTKTAFDLILLEKAETSKLTNPSLPDTRACRIPAFHGNVYVKVCHACCVLNQGGKR
	NTVLPFGPASTGKSIIAQALAQAVNGVCTNAANVFPFNDCKNKLIVVEEAGNFGQ
	QVNFKAICSGQTIKIDQKSGSKQIEPTFVINTNENITVVRIGCEERPEHTOPIRD
	RMLNIHLTHLPDGLVDKNWPMICAWLVKNGYQSTWASYSYKAKGVKPDWENAE
	PKVPTPINLIGSARSPTTPKSTPLSONVALTPLASDLELALPEWSTPNTPVAGTAE
	TQNTGEGSKACQDQGLSPTWSEIEEDLRACFGEPLKXDFSEPLND"

ORIGIN

Query Match	99.8%;	Score	2015.8;	DB	6;	Length	2019;
Best Local Similarity	99.9%;	Pred. No.	0;				
Matches	2017;	Conservative	0;	Mismatches	2;	Indels	0;
	Gaps						
Qy	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60				
Db	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60				
Qy	61	AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTTCAACTGAATGGAAAA	120				
Db	61	AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTTCAACTGAATGGAAAA	120				
Qy	121	GATATCGGATGGAAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	180				
Db	121	GATATCGGATGGAAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	180				
Qy	181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACACAGTGGAT	240				
Db	181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACACAGTGGAT	240				
Qy	241	GAAATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAAAATAATGTTTATTGAA	300				
Db	241	GAAATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAAAATAATGTTTATTGAA	300				
Qy	301	GTGCTTTACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCAACATGATGG	360				
Db	301	GTGCTTTACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCAACATGATGG	360				
Qy	361	GGAAAGACCAAGGCTGGCACTGCGCATGTACTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420				
Db	361	GGAAAGACCAAGGCTGGCACTGCGCATGTACTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420				
Qy	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC	480				
Db	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC	480				
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT	540				
Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT	540				
Qy	541	GAGTGGGTTACTCTACTTCTTATAGCATAGCAAAACCAAAAAGACTATACCAAGTGT	600				
Db	541	GAGTGGGTTACTCTACTTCTTATAGCATAGCAAAACCAAAAAGACTATACCAAGTGT	600				
Qy	601	GTTCTTTTGGAAAACATGATTGCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT	660				
Db	601	GTTCTTTTGGAAAACATGATTGCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT	660				

```
QY 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGGAAACTAACTTTTAA 720
Db 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGGAAACTAACTTTTAA 720
QY 721 AAAGAAGGCGAGCGCCATCTAGTGAAGAACTATACACTGATGACATCGGCGCAGAAACG 780
Db 721 AAAGAAGGCGAGCGCCATCTAGTGAAGAACTATACACTGATGACATCGGCGCAGAAACG 780
QY 781 GTTGAACACACAGTAACCACTGCGCAGGAACTTAAGCGCGCAGAACTCAAACTAAAAA 840
Db 781 GTTGAACACACAGTAACCACTGCGCAGGAACTTAAGCGCGCAGAACTCAAACTAAAAA 840
QY 841 GAAAGTTTCTATTAAAACTACACTTAAGAGCTGGTGCATAAAAAGAGCTAACCTCACCAGAG 900
Db 841 GAAAGTTTCTATTAAAACTACACTTAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900
QY 901 GACTGGATGATGATGACGACGACAGCTTACATTGAAATGATGGCTCAACAGAGTGAGAA 960
Db 901 GACTGGATGATGATGACGACGACAGCTTACATTGAAATGATGGCTCAACAGAGTGAGAA 960
QY 961 AACCTGCTGAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db 961 AACCTGCTGAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAATTTAGAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1021 TTTGACTTAATTTAGAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
QY 1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGGTGAACATATGTTAAAGTTTGCCTAGCT 1140
Db 1081 ACAAGAGCTGCAAGATTTTGTCTTTTATGCTGGTGAACATATGTTAAAGTTTGCCTAGCT 1140
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAGAAATACCTGTTTATTTATCATGACCA 1200
Db 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAGAAATACCTGTTTATTTATCATGACCA 1200
QY 1201 GCCAGACAGCGCAATCTATTATGCAAGCCATAGCAACAAGCAGTTGGCAATGTTGGT 1260
Db 1201 GCCAGACAGCGCAATCTATTATGCAAGCCATAGCAACAAGCAGTTGGCAATGTTGGT 1260
QY 1261 TGCTATAATGACGCAATGTAACCTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGCTATAATGACGCAATGTAACCTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
QY 1321 TGGGTAGAGAGAGCTGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCATTGCTCT 1380
Db 1321 TGGGTAGAGAGAGCTGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCATTGCTCT 1380
QY 1381 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACCA 1440
Db 1381 GGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACCA 1440
QY 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGACAGAGATGCTTAAACATTCTAACATCATCATACATCTTGCCT 1560
Db 1501 GAACACACTCAACCAATCAGACAGAGATGCTTAAACATTCTAACATCATCATACATCTTGCCT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCCATGATTGCTGTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCCATGATTGCTGTTGGTAAAG 1620
QY 1621 AATGGTTACCAATCTACCTAGCGCAGTACTGTGCTAAATGGGCAAGTTCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACCTAGCGCAGTACTGTGCTAAATGGGCAAGTTCTGATTGG 1680
QY 1681 TCAGAAACTGGGCGGAGCCAAAGGTGCCAACTCCCTATAAATTTACTAGGTTGGGACGC 1740
Db 1681 TCAGAAACTGGGCGGAGCCAAAGGTGCCAACTCCCTATAAATTTACTAGGTTGGGACGC 1740
```

```
QY 1741 TCACCATTCACACACCCGAAAAAGTAGCGCTCTCAGCAGAACTATGCACTAACTCCACTT 1800
Db 1741 TCACCATTCACACACCCGAAAAAGTAGCGCTCTCAGCAGAACTATGCACTAACTCCACTT 1800
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATAACTCTCTGTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATAACTCTCTGTTGG 1860
QY 1861 GGCACCTGCGAGAAACCCAGAAACACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db 1861 GGCACCTGCGAGAAACCCAGAAACACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
QY 1921 CTGAGCCCCAACTTGCTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
Db 1921 CTGAGCCCCAACTTGCTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019

RESULT 7
AX137747
LOCUS AX137747 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 12 from Patent EP1077260.
ACCESSION AX137747
VERSION AX137747.1 GI:14273921
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1. Nueesch, J. and Rommelaere, J.
Parvovirus ns1 variants
Patent: EP 1077260-A 12 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="Cac39995.1"
/db_xref="GI:14273922"
/translation="MAGNAYSDEVLGATNWLKESNOEVSFVEKNENVOLNGKDIGW
NSYKXELQDEDELSLQRGAEITWDSSEWENETTVDEMTKKQVIFDSLKKCLFEVL
NTKNTPPGDVNWFQHWGKQDQWHCHVLIGKDFSAQGWKWRRLQNLVWGRWLVT
CNVQLTPABRIKLREIAEDNEWVTLTKYKQTKDYTKCVLFGNMIAFYFLTKKIS
TSPRDGGVFLSSDGSWKTNFLKEGRLVSKLYTDMRPETVETVTTTAQETKRGRI
OTPKVSIKTTLLKELVHKVTSPEDEMMQPSYIEMMAQPGGKLNKLTLEICTLTL
ARTKAPDLLEKAEISKLITNPSLPTETCRIFAPHWYVKVCHAI CVLINRQGGKR
NAVLFPGASGTIIQAIQAAGVNGVCYNAANVPNDCTNKNLIWVEAGNFGQ
QYNQKAICTGIRIDQKGSQIETPTVIMTNENITVVRIGCEEPHTQPIRD
RMNLTHLTLPGLVDKNWPMICAWLVKNGYQSTWASYCAKWGKVPDWSNWA
PKVPTPINLLGSARSPFTPKSTPSQNTALTPLASDLEDLALPEWSTPTNPVAGTAE
TQNTGBAGSKACQDGLSPTWSEIBEDLRACFGAEPLKKDFSEPLND"

Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60
Db 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60
QY 61 AGTAACACAGAGAGTGTCTCAITTTGTTTTTAAAAATGAAAAATGTTTCAACTGAATGAAAA 120
Db 61 AGTAACACAGAGAGTGTCTCAITTTGTTTTTAAAAATGAAAAATGTTTCAACTGAATGAAAA 120
```

QY	121	GATATCGGATGGATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA	180
Db	121	GATATCGGATGGATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA	180
QY	181	CGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGAATGGGAAACCAACAGTGGAT	240
Db	181	CGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGAATGGGAAACCAACAGTGGAT	240
QY	241	GAATGACCAAAAAAGCAAGTATTCATTTTGGTCTTTGTTTAAATAATGTTTATTGAA	300
Db	241	GAATGACCAAAAAAGCAAGTATTCATTTTGGTCTTTGTTTAAATAATGTTTATTGAA	300
QY	301	GTGCTTAAACCAAGAAATATTTCTGCTGATGTTAAATTTGTTTAAATAATGTTTATTGAA	360
Db	301	GTGCTTAAACCAAGAAATATTTCTGCTGATGTTAAATTTGTTTAAATAATGTTTATTGAA	360
QY	361	GGAAAGACCAAGCTGCGACTGCGATGTAATTTGGAGGAAAGGACTTTTGTCAAGCT	420
Db	361	GGAAAGACCAAGCTGCGACTGCGATGTAATTTGGAGGAAAGGACTTTTGTCAAGCT	420
QY	421	CAAGGAAATGGTGGAGGCAAACTAAATTTTACTGAGCAGATGTTTGGTAACAGCC	480
Db	421	CAAGGAAATGGTGGAGGCAAACTAAATTTTACTGAGCAGATGTTTGGTAACAGCC	480
QY	481	TGTAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
Db	481	TGTAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
QY	541	GAGTGGTGTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGTATACCAAGTGT	600
Db	541	GAGTGGTGTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGTATACCAAGTGT	600
QY	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660
Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660
QY	661	CCACCAAGAGACGAGGCTATTTTCTAGCAGTGAATCTGGCTGGAAACATACTTTTAA	720
Db	661	CCACCAAGAGACGAGGCTATTTTCTAGCAGTGAATCTGGCTGGAAACATACTTTTAA	720
QY	721	AAAGAGGCGGAGCGCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG	780
Db	721	AAAGAGGCGGAGCGCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG	780
QY	781	GTTCGAAACCAAGTAACTGCGGAGGAACTAAAGCGGCGAGAAATTCAAAATAAAAAA	840
Db	781	GTTCGAAACCAAGTAACTGCGGAGGAACTAAAGCGGCGAGAAATTCAAAATAAAAAA	840
QY	841	GAAGTTTCTAATTAATACTACATTAAGAGCTGGTGCATTAAGAGAGTAACTCACCAGAG	900
Db	841	GAAGTTTCTAATTAATACTACATTAAGAGCTGGTGCATTAAGAGAGTAACTCACCAGAG	900
QY	901	GACTGGATGATGACGAGCAGCAGTATGTAATGATGCTCAACAGGTGAGAA	960
Db	901	GACTGGATGATGACGAGCAGCAGTATGTAATGATGCTCAACAGGTGAGAA	960
QY	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCAGAAACCAAAACAGCA	1020
Db	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCAGAAACCAAAACAGCA	1020
QY	1021	TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCGCTGAC	1080
Db	1021	TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCGCTGAC	1080
QY	1081	ACAAGAACCTGCAAGATTTTGTCTTTCATGGCTGGAATATGTTTAAAGTTTCCCATGCT	1140
Db	1081	ACAAGAACCTGCAAGATTTTGTCTTTCATGGCTGGAATATGTTTAAAGTTTCCCATGCT	1140
QY	1141	ATTGCTGTGTTTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTATTTATGACCA	1200
Db	1141	ATTGCTGTGTTTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTATTTATGACCA	1200

QY	1201	GCAGCACAGGCAAACTATTATTGCACAGCCATAGCAACAAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCAGCACAGGCAAACTATTATTGCACAGCCATAGCAACAAGCAGTTGGCAATGTTGGT	1260
QY	1261	TGCTATAATGCGACCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAAGAACTTCATT	1320
Db	1261	TGCTATAATGCGACCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAAGAACTTCATT	1320
QY	1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCATTTGCTCT	1380
Db	1321	TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCATTTGCTCT	1380
QY	1381	GGTCAAGCTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAACAGATTGAAACCAACCA	1440
Db	1381	GGTCAAGCTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAACAGATTGAAACCAACCA	1440
QY	1441	GTCAATCATGACCAAAATGAGAACATTCAGTGGTCAAGTATAGCTGCGAAGAAAGACCA	1500
Db	1441	GTCAATCATGACCAAAATGAGAACATTCAGTGGTCAAGTATAGCTGCGAAGAAAGACCA	1500
QY	1501	GAAACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTCATCTAAACATACCTTGCCT	1560
Db	1501	GAAACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTCATCTAAACATACCTTGCCT	1560
QY	1561	GGTGACTTTGGTTGGTTGTCACAAAATGAATGCGCCCATGATTTTGTGCTTGGTTGAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGTCACAAAATGAATGCGCCCATGATTTTGTGCTTGGTTGAAG	1620
QY	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
Db	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
QY	1681	TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTTGGCAGCG	1740
Db	1681	TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTTGGCAGCG	1740
QY	1741	TCACCAATTCAGACACCGAAAGTAGCGCTCTCAGCAGAACTATGCACTAACTCCACTT	1800
Db	1741	TCACCAATTCAGACACCGAAAGTAGCGCTCTCAGCAGAACTATGCACTAACTCCACTT	1800
QY	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATACTCTCTGTTGCG	1860
Db	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATACTCTCTGTTGCG	1860
QY	1861	GGCACTGCGAGAAACCCAGAACACTGGGGAGCTGGTTTCCAAAGCTGCCAAGATGTTCAA	1920
Db	1861	GGCACTGCGAGAAACCCAGAACACTGGGGAGCTGGTTTCCAAAGCTGCCAAGATGTTCAA	1920
QY	1921	CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG	1980
Db	1921	CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG	1980
QY	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019	
Db	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019	

RESULT 8  
MWICG 5085 bp ss-DNA linear VRL 05-JUL-1995  
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM, complete genome.  
DEFINITION complete genome.  
ACCESSION M12032  
VERSION M12032.1 GI:332289  
KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.  
SOURCE Mouse minute virus  
ORGANISM Mouse minute virus  
REFERENCE 1 (bases 1 to 5085)  
AUTHORS Astell,C.R., Gardner,E.M. and Tattersall,P.  
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic



ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948; Conservative	0; Mismatches 71; Indels 0; Gaps 0;	
QY	1 ATGCTGGAAATCCTTACTCTGATGAGTTTTCGGGACCAACCACTGGTTAAAGGAAAA	60
DB	262 ATGGCTGGAAATCCTTACTCTGATGAGTTTTCGGGACCAACCACTGGTTAAAGGAAAA	321
QY	61 AGTAACCAAGAGTGTCTCATTTGTTTAAAAATGAAATGTCTCAACTGAATGCAAAA	120
DB	322 AGTAACCAAGAGTGTCTCATTTGTTTAAAACTGAGATGTCTCAACTGAATGCAAAA	381
QY	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGAGTGAATCTTTACAA	180
DB	382 GATATCGGATGGAATTAATACAAAAAGAGCTCGAGGAGCAGAGTGAATCTTTACAA	441
QY	181 CGAGGACGGAATCTTCTGGACCAAGAGCGAGACATGGAATGGGAAACCAAGTGGAT	240
DB	442 CGAGGACGGAATCTTCTGGACCAAGAGCGAGACATGGAATGGGAAATCTACAGTGGAT	501
QY	241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA	300
DB	502 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA	561
QY	301 GTGCTTAAACAAGAAATATTTCTCGTGATGTTTAATTTGGTTTGTGCAACATGAATGG	360
DB	562 GTGCTTAAACAAGAAATATTTCTCGTGATGTTTAATTTGGTTTGTGCAACATGAATGG	621
QY	361 GGAAGACCAAGCTGGCACTGGCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT	420
DB	622 GGAAGACCAAGCTGGCACTGGCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT	681
QY	421 CRAAGGAAATGGTGGAGAGCACTAAATGTTTACTTGGAGCAGATGTTGTTAAACAGCC	480
DB	682 CRAAGGAAATGGTGGAGAGCAGCTAAATGTTTACTTGGAGCAGATGTTGTTAAACAGCC	741
QY	481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540
DB	742 TGTAAATGTGCACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAGT	801
QY	541 GAGTGGTTACTCTTACTTATTAAGCATAGCAAAACCAAAAGAGCTATACCAAGTGT	600
DB	802 GAGTGGTTACTCTTACTTATTAAGCATAGCAAAACCAAAAGAGCTATACCAAGTGT	861
QY	601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
DB	862 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	921
QY	661 CCACCAAGAGACGAGGCTATTTTCTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
DB	922 CCACCAAGAGACGAGGCTATTTTCTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	981
QY	721 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGAGAAACG	780
DB	982 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGAGAAACG	1041
QY	781 GTTGAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATCAAACTAAAAA	840
DB	1042 GTTGAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATCAAACTAAAAA	1101
QY	841 GAAGTTTCTATTAACCACTTAAAGAGCTGGTGATTAAGAGAGTAACCTCACAGAG	900
DB	1102 GAAGTTTCTATTAACCACTTAAAGAGCTGGTGATTAAGAGAGTAACCTCACAGAA	1161
QY	901 GACTGGATGATGACGAGCAGTATACATTTGAAATGATGCTCAACAGGTGGAGAA	960
DB	1162 GACTGGATGATGACGAGCAGTATACATTTGAAATGATGCTCAACAGGTGGAGAA	1221
QY	961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAAACCAAAACAGCA	1020
DB	1222 AACTGCTGAAAAATACGCTAGAGATTTGTACGCTAATCTAGCCAGAAACCAAAACAGCA	1281

QY	1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTTCACCTGCTGAC	1080
DB	1282 TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTTCACCTGCTGAC	1341
QY	1081 ACAAGAACCTGCAGAAATTTTTCCTTTTCATGGCTGGAATATATGTTTAAAGTTTGCCTGCT	1140
DB	1342 ACAGAACCTGCAGAAATTTTTCCTTTTCATGGCTGGAATATATGTTTAAAGTTTGCCTGCT	1401
QY	1141 ATTTGCTGTGTTTTTAAACAGACAAAGAGGCAAAAAGAAATACATGTTTATTTATGACCA	1200
DB	1402 ATTTGCTGTGTTTTTAAACAGACAAAGAGGCAAAAAGAAATACATGTTTATTTATGACCA	1461
QY	1201 GCCAGACAGGCAAAATCTATTTATGCAAGCCATAGCAAGACAGTGTGGCAATGTTGCT	1260
DB	1462 GCCAGTACAGGCAAAATCTATTTATGCAAGCCATAGCAAGACAGTGTGGTAAATGTTGCT	1521
QY	1261 TGTATAATGCAAGCAATGTAACCTTTTCCATTTTAATGACTGTACCAACCAAGAACTTTGATT	1320
DB	1522 TGTATAATGCAAGCAATGTAACCTTTTCCATTTTAATGACTGTACCAACCAAGAACTTTGATT	1581
QY	1321 TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT	1380
DB	1582 TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT	1641
QY	1381 GGTCAAGCTATTTCGCATTTGATCAAAAAGGAAAGGAGCAAAACAGATTTGAAACCAACCA	1440
DB	1642 GGTCAAACTATTTCGCATTTGATCAAAAAGGAAAGGAGCAAAACAAATTTGAAACCAACCA	1701
QY	1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA	1500
DB	1702 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA	1761
QY	1501 GAACACATCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACATTCGCT	1560
DB	1762 GAACACATCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACATTCGCT	1821
QY	1561 GGTGACTTTGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGCTTGGTTGTTAAAG	1620
DB	1822 GGTGACTTTGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGCTTGGTTGTTAAAG	1681
QY	1621 AATGGTTACCAATCTACCATGGCAAGTACTGCTTAAATGGGGCAAAAGTTCTCTGATTGG	1680
DB	1882 AATGGTTACCAATCTACCATGGCAAGTACTGCTTAAATGGGGCAAAAGTTCTCTGATTGG	1941
QY	1681 TCAGAAAACTGGGCGGAGCAAAAGTGCACCTCTCTATAAAATTTACTAGTTTCGGCACGC	1740
DB	1942 TCAGAAAACTGGGCGGAGCAAAAGTGCACCTCTCTATAAAATTTACTAGTTTCGGCACGC	2001
QY	1741 TCACCAATTCAGACACCCGAAAAAGTACGCTCTCAGCCAGAACTATGACATTAACCTT	1800
DB	2002 TCACCAATTCAGACACCCGAAAAAGTACGCTCTCAGCCAGAACTATGCAATAACTCCACTT	2061
QY	1801 GCATCGATCTCGAGACCTGGCTTTAGAGCTTTAGAGCAGACCAAAATCTCTCTGTTGCG	1860
DB	2062 GCATCGATCTCGAGACCTGGCTTTAGAGCTTTAGAGCAGACCAAAATCTCTCTGTTGCG	2121
QY	1861 GGCATCTGAGAAACCCAGAAACACTGGGGAGAGTGGTTTCCAAAGCTGCCAAGATGTTCAA	1920
DB	2122 GGCATCTGAGAAACCCAGAAACACTGGGGAGAGTGGTTTCCAAAGCTGCCAAGATGTTCAA	2181
QY	1921 CTGAGCCCAACTTTGGTTCAGAGATTCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAACCG	1980
DB	2182 CTGAGCCCAACTTTGGTTCAGAGATTCGAGGAGGATTTGAGAGCGTGTTCGGTGGGAACCG	2241
QY	1981 TTGAAAGAAAGACTTCAGCGAGCGCTGAACTGGACTAA	2019
DB	2242 TTGAAAGAAAGACTTCAGCGAGCGCTGAACTGGACTAA	2280

RESULT 9  
MOU34253  
LOCUS

MOU34253 4764 bp DNA linear VRL 21-AUG-1996

[illegible]

Db 1802 ACAGAAACTGGCGGAGCCGAGGTCCTTATTAATCACTAGGTCGGCAGCG 1861  
Qy 1741 TCACCATTCAGACACCGAAAGTACGCTCTCAGCAGAACATATGCACTAATCCACTT 1800  
Db 1862 TCACCATTCAGACACCGAAAGTACGCTCTCAGCAGAACATATGCACTAATCCACTT 1921  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGGACACACCAATATCTCCTGTTGCG 1860  
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGGACACACCAATATCTCCTGTTGCG 1981  
Qy 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Db 1982 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGGTCAA 2041  
Qy 1921 CTGAGCCCAACTGGTTCAGAGATCGAGAGATTGAGAGCGTCTTCGGTGGGGAACCG 1980  
Db 2042 CTGAGCCCAACTGGTTCAGAGATCGAGAGATTGAGAGCGTCTTCGGTGGGGAACCG 2101  
Qy 1981 TTGAAGAAGACTTCAGGAGCGCTCAACTTGGACTAA 2019  
Db 2102 TTGAAGAAGACTTCAGGAGCGCTCAACTTGGACTAA 2140

RESULT 10  
PAMVMI  
LOCUS PAMVMI 5087 bp DNA linear VRL 01-JUL-1999  
DEFINITION Mouse parvovirus minute virus immunosuppressive variant genome (= PAMVMI).  
ACCESSION X02481  
VERSION X02481  
KEYWORDS coat protein; genome; origin of replication; overlapping genes; terminal repeat; unidentified reading frame.  
SOURCE Mice minute virus  
ORGANISM Mice minute virus  
REFERENCE 1 (bases 1 to 5087)  
AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.  
TITLE DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice  
JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)  
MEDLINE 85242059  
PUBMED 3855242  
COMMENT For the fibroblast-specific strain (MVmp) sequence see <PAMV2>. The genomes of MVmp and MVmi (immunosuppressive variant) have more than 96% of their sequence in common.  
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES  
source 1..5087  
/organism="Mice minute virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10794"  
misc\_feature 1..116  
/notes="terminal hairpin region"  
promoter 175..183  
/note="TATA box"  
gene 2405..4558  
/gene="VP1"  
CDS <2405..4558  
/gene="vp1"  
/function="coat protein"  
/codon\_start=1  
/product="vp1 protein"  
/protein\_id="CAB46507.1"  
/db\_xref="GI:5419928"  
/db\_xref="GOA:P07302"  
/db\_xref="Swiss-Prot:P07302"  
/translation="VPPGYLPGNSLDGSEPTNPSDAAKEHDEAYDQYIKSGRNP YLYFSADQRFIDQTKADMGKGVHYFFRTKRAFAPKLADSEFPGTSGVRAGRT RPAYTIPINQARAKLTSSAAQSSQSDGTSGDGGNGVHSAARVERAADPGGS GGGGGGGVGTSGSYDNOYHFLDGMVEITATLRLVHLMPSKSENYCRIRVHN TTDTSVKGNWAKDDAEQIWTPLSLVDANAGWLQPSDWQYICNTMSQLNLSLDOE IFNVVLKTVTEQSGGAQIKIYNNDLTACMVAVDNSNLLPYTPAANSMETLGFYFWK FTIASPYRYFYFCVDRDLDSVTYENQEGTIEHNVMGTPKGMNSQFFTIENQIITLLRTG

DEPATCTYFDTNPKVLTHWTOTNROLGPPILLSTRPADTDTAGTLTAQSSRBGATOM  
EUNVNSEAINTRAQVGFQCPHNDFEASRAGPFAAPKVADVTQGVDRREANGSVISY  
GKHGHNWAHAGPAPERYTWDETFNFGSDRDRGFIQSAFLVPPPLUNGILLTNANPG  
TKNDIHFSNVFNSYBELTAPSHSPYVQGIWDLKDLHSHKPLHITAPFVCKNNAP  
GQMLVRLGNLTDQYDPNGATLSRIYVTGTFWKGKLTMRKLRANTTNPNVYQVSVE  
DNGNSYMSVTWKLPATATGNMQSVPLITRPVARNY"  
2792..4558  
/gene="VP2"  
CDS <2792..4558  
/gene="VP2"  
/function="coat protein"  
/codon\_start=1  
/product="VP2 protein"  
/protein\_id="CAB46508.1"  
/db\_xref="GI:5419929"  
/db\_xref="GOA:P07302"  
/db\_xref="Swiss-Prot:P07302"  
/translation="TMSDGTSGDGGNGVHSAARVERAADPGSGGGGGGGGVGS  
TGSYDNOYHFLDGMVEITATLRLVHLMPSKSENYCRIRVHNTTDTSVKGNWAKD  
DAHEQIWTPLSLVDANAGWLQPSDWQYICNTMSQLNLSLDOEIPNVVLKTVTEQD  
SGGAQIKIYNNDLTACMVAVDNSNLLPYTPAANSMETLGFYFWKPLHITAPFVCK  
DRDLSTYENQEGTIEHNVMGTPKGMNSQFFTIENQIITLLRTGDEFTATGYTFTDN  
PVKLTHTWOTNROLGPPILLSTRPADTDTAGTLTAQSSRBGATOMVNVSEAIRTRP  
AOKVFCQPHNDFEASRAGPFAAPKVADVTQGVDRREANGSVISYKQKHGHNWAHGP  
APERIYTWDETFNFGSDRDRGFIQSAFLVPPPLUNGILLTNANPGTENDIHFSNVFS  
YGLTAFSPSPYVQGIWDLKDLHSHKPLHITAPFVCKNNAPGQMLVRLGNLTD  
QYDPNGATLSRIYVTGTFWKGKLTMRKLRANTTNPNVYQVSVEDNGNSYMSVTWK  
LPATGNMQSVPLITRPVARNY"  
4603..4608  
misc\_feature  
/note="polyadenylation signal"  
4821..4826  
misc\_feature  
/note="polyadenylation signal"  
4878..5087  
misc\_feature  
/note="terminal hairpin region"

ORIGIN

Query Match 94.2%; Score 1902.2; DB 14; Length 5087;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCACTGGTTAAAGGAAAA 321  
Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTAAAAAATGAAATGTTCAACTGAATGGAATA 120  
Db 322 AGTAACCAAGAGTGTCTCATTTGTTTAAAAAATGAGGATGTTCAACTGAATGGAATA 381  
Qy 121 GATATCGGATGGATAGTTTACAAAAAAGAGCTCCAGAGGACGAGCTGAATCTTTACAA 180  
Db 382 GATATCGGATGGAAATAATTACAAAAAGGAGCTCGAGGAGACGAGCTGAAATCTTTACAA 441  
Qy 181 CGAGGAGCGAAACTACTTTGGGACCAAGAGGACATGGAATGGGAAACACAGTGGAT 240  
Db 442 CGAGGAGCGGAAACTACTTTGGGACCAAGAGGACATGGAATGGGAAATCTACAGTGGAT 501  
Qy 241 GAAATGACCAAAAAGCAAGTATTTCATTTTGGTTTCTTTGTTTAAAAAATGTTTATTGAA 300  
Db 502 GAAATGACCAAAAAGCAAGTATTTCATTTTGGTTTCTTTAGTCTTTTAAATAATGTTTGTGAA 561  
Qy 301 GTGCTTAAACAAAGAATATATTTCTCTGGTATGTTAATTTGGTTTGTGCAACATGAATGG 360  
Db 562 GTGCTTAAACAAAGAATATATTTCTCTGGTATGTTAATTTGGTTTGTGCAACATGAATGG 621  
Qy 361 GGAAGAGACCAAGGCTGGCACTGCCCATGTACTAATTTGGAGAAAGAGCTTTTAGTCAAGCT 420  
Db 622 GGAAGAGACCAAGGCTGGCACTGCCCATGTACTAATTTGGAGGCAAGAGCTTTTAGTCAAGCT 681  
Qy 421 CAAGGAAATGTTGGGAAAGGCAACTAAATTTTACTTGGAGCAGATGTTTGGTGAACAGCC 480  
Db 682 CAAGGAAATGTTGGGAAAGGCAACTAAATTTTACTTGGAGCAGATGTTTGGTGAACAGCC 741  
Qy 481 TGTAATGTGCAACTAACACACGAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGCAAT 540







```
OTKKEYSIKTLKELVHKSIVTSPEDMMWQPOSYIEWMAQPGGENIKNTLEICTLTLL
ARTKTAFDLLEKAEKTSKINFLSPDTRCKIPAFHGWNYIKVCHALCCVLNRQGGKR
NTVLFPHGASTGKSIIAQALQAGVNGVCYNAANVFPNFDCTNKNKLIWVEEAGNFGQ
OWNFQKISGOTIRIDOKGSKQIEPTVIMTNENITVVKICEERPEHTPIRD
RMLNHLTLPLGDFLVDKNQWPMICAMLVKNGYQSTWASCYAKWGVKPDWTENWAE
PKVPTINSLSARSPTTPKSPPLSNVALTPLASDLDELDALEPMWSTENTPVAGTAE
TQNTGKSKACODGQSLSPWBSIEDLRACFGAEPLKXDFSEPLND"
join(2290..2320,2403..4561)
/genes="Vp1"
join(2290..2320,2403..4561)
/genes="Vp1"
/codon_start=1
/product="capsid protein"
/protein_id="AAA61406.1"
/db_xref="GI:525327"
/translation="MAPPAKRAKRWPPGYKYLPGNSLDQSEPTNPSPAARKEHDE
AYDKYIKSGNVPFLYFSADQRIDIDOKADGWGVHVFYFTRKFAFAPRLASSEPE
GTSVSIAGRTKPPPAHIFINOARAKKRSLAAQOORTOTMSDAGSPQSGSAVQAA
RVERAADGPGGSGGGGGVGVSGSYDNQTHYRFLSDGWVEITAYSTRVHLMNP
KSENYCRVRHNTNTRTAGNMAKDAHEQIWTPLSLIDSNWGVWFQPSDWOFICNN
MSHVNLSLQDELFNVIKTVTEQNTGAEAVKIYNNDLTASMMVALDSNNILPFTPT
DNETLGFYFWKFTSPRYFYENCDSLSVTYDQTSIIVDTMANASGLSOFYFIE
NTQRIOLLRTGDEFATGYFETPEIKLSHTWOSNRQLGPOPOITLPLADNENATLV
TRGDRSGIQTQISGNDVTEATVRPAQVGFPCPDHNFETSRAGFPVVPVNAVTOGN
EHDANGSLRTYDKQGDWGSNNKERTFWDALSYDSGRWADRCFINATPFTSPPAL
NLTITNSDPIGNKTAIHYQNVFNSYGLTAFPHPAPIYPOGQIWDKELDLDEHPRUHA
QAPFVCKNNAAGQLLVRLAPNLTDQDPSNSTLSRVTYGTGTFWFKGKJLTKLRPNA
TWNPVTVQVSAQYQNENEYMSIHKWLPATATGNMQSIPLLSRPVARNTY"
```

## ORIGIN

```
Query Match          94.1%; Score 1899; DB 14; Length 5144;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATGCTGGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTTAAAGGAAAA 60
DB 265 ATGCTGGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTTAAAGGAAAA 324

QY 61 AGTAACCAAGAGTGTCTCATTTGTTTTTAAATAAGTGTCAACTGAAATGGAATA 120
DB 325 AGTAACCAAGAGTGTCTCATTTGTTTTTAAATAAGTGTCAACTGAAATGGAATA 384

QY 121 GATATCGGATGGAATAGTTTACAAATAAGCTGAGAGACGAGCTGAAATCTTTACAA 180
DB 385 GATATCGGATGGAATAGTTTACAAATAAGCTGAGAGACGAGCTGAAATCTTTACAA 444

QY 181 CGAGGACGGAATACTCTGGGCCCAAGCAGACATGGAATGGGAACCAACAGTGGAT 240
DB 445 CGAGGACGGAATACTCTGGGCCCAAGCAGACATGGAATGGGAATCTACAGTGGAT 504

QY 241 GAAATGACCAAAAAGCAATGATTTTCAATTTTGGTTTAAATAAGTGTTTATTTGAA 300
DB 505 GAAATGACCAAAAAGCAATGATTTTCAATTTTGGTTTAAATAAGTGTTTATTTGAA 564

QY 301 GTGCTTTAAACAAAGAATAATATTTCTGGTGATGTTAAATTTGGTTTGGCAACATGATGG 360
DB 565 GTGCTTAGCAACAAAGAATAATAGCTCTGCTGATGTTACTTGTGTTGGCAGCATGATGG 624

QY 361 GGAAGAAGCAAGCTGGCAGTGCATGTAATTTGGAGAGAAAGGACTTTAGTCAAGCT 420
DB 625 GGAAGAAGCAAGCTGGCAGTGCATGTAATTTGGAGAGAAAGGACTTTAGTCAAGCT 684

QY 421 CAAGGGAATTTGGTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC 480
DB 685 CAAGGGAATTTGGTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC 744

QY 481 TGTAAATGTCAACTTAACACAGCTGAAAGAATAATTAACCTAAGAGAAATAGCAGAGCAAT 540
DB 745 TGTAAATGTCAACTTAACACAGCTGAAAGAATAATTAACCTAAGAGAAATAGCAGAGCAAT 804

QY 541 GAGTGGGTACTTACTTACTTATTAAGCATTAAGCAAAACCAAAAGACTATACCAAGTGT 600
DB 805 GAGTGGGTACTTACTTACTTATTAAGCATTAAGCAAAACCAAAAGACTATACCAAGTGT 864
```

```
QY 1681 TCAGAAACTGGCGGAGCCAAAGTCCCACTCTATAAATTTACTAGTTCGGCAGC 1740
Db 1945 ACAGAAACTGGCGGAGCCAAAGTCCCACTCTATAAATTTACTAGTTCGGCAGC 2004
QY 1741 TCACCAATTCAGACACCGAAGTACGCTCTCAGCCAGCACTATGCACTTAATCCACTT 1800
Db 2005 TCACCAATTCAGACACCGAAGTACGCTCTCAGCCAGCACTATGCACTTAATCCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCTGTTCG 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCTGTTCG 2124
QY 1861 GGCACTCGAAGAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGCTCAA 1920
Db 2125 GGCACTCGAAGAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGCTCAA 2184
QY 1921 CTGAGCCCAACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACG 1980
Db 2185 CTGAGCCCAACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACG 2244
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2019
Db 2245 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2283

RESULT 12
MOU34254
LOCUS MOU34254 4764 bp DNA linear VRL 21-AUG-1996
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE Mouse parvovirus 1c
ORGANISM Mouse parvovirus 1c
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
J Gen Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 9620434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1. .4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 122 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 181
QY 61 AGTAACAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAATA 120
Db 182 AGTAACAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAATA 241
QY 121 GATATCGGATGGAATAGTTTCAAAAAGAGCTCGCAGGACGAGCTGGAATCTTTTACA 180
Db 242 GATATCGGATGGAATAGTTTCAAAAAGAGCTCGCAGGACGAGCTGGAATCTTTTACA 301
QY 181 CGAGGAGCGGAAACTTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240
```

```
Db 302 CGAGGAGCGGAAACTTACTTGGGACCAAGCGAGGACATGGAATGGAATCTACAGTGGAT 361
QY 241 GAAATCACCAAAAGCAAGTATTTCATTTTGGTTAAAGAAATGTTTATTATTGAA 300
Db 362 GAAGTACCAAAAGCAAGTATTTCATTTAGTACTCTTTAGTTAAAGAAATGTTTATTGAA 421
QY 301 GTGCTTAAACACAAAGAAATATATTTCTCTGGTGTATTTAAATTTGGTTTGTCAACATGAATGG 360
Db 422 GTGCTTAAACACAAAGAAATATATCTCTGGTGTATTTAAATTTGGTGTGAGCATGAATGG 481
QY 361 GHAAGAACAAGCTGGCACTGCGCATGTACTTAATTTGGAGGAAGAGACTTTAGTCAAGCT 420
Db 482 GGAAGAGACCAAGCTGGCACTGCGCATGTACTTAATTTGGAGGAAGAGACTTTAGTCAAGCT 541
QY 421 CAAGGAAATGTTGGAGGAAGCAACTAAATTTTACTTGGAGCAGATGTTTGTGTAAACAGCC 480
Db 542 CAAGGAAATGTTGGAGGAAGCAAGCTAAAGTTTACTTGGAGCAGATGTTTGTGTAAACAGCC 601
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGAACAT 540
Db 602 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGACAGT 661
QY 541 GAGTGGTTACTCTACTTACTTATAGCATATAGCAAAACCAAAAGAACTATATACCAAGTGT 600
Db 662 GAATGGTTTACCTTGTCTCTTATAGCATATAGCAAAACCAAAAGAACTATATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATTCCTTACTATTTTAACTAAAGAAATATAAGCACTAGT 660
Db 722 GTTCTTTTGGAAACATGATTCCTTACTATTTTAACTAAAGAAATATAAGCACTAGT 781
QY 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGAATCTGGCTGGAAAACTTAACCTTTTAA 720
Db 782 CCGCCAAAGGACGAGGCTATTTCTTAGTGAATCTGGCTGGAAAACTTAACCTTTTAA 841
QY 721 AAGAAGGCGAGGCGCATCTAGTGAGCAAACTATATACACTGTATGACATGCGGCCAGAAACG 780
Db 842 AAGAAGGCGAGGCGCATCTAGTGAGCAAACTATATACACTGTATGACATGCGGCCAGAAACG 901
QY 781 GTTGAACACACAGTAAACCACTGCGCAGGAAACTTAAGCGCGGCAAGATTCAAACTTAAATA 840
Db 902 GTTGAACACACAGTAAACCACTGCGCAGGAAACTTAAGCGCGGCAAGATTCAAACTTAAATA 961
QY 841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACCTCACAGAG 900
Db 962 GAGTTTCTATTAAAAACCACTTAAAGAGCTAGTGCATATAAGAGAGTAACTTCAACAGAA 1021
QY 901 GACTGATGATGATGCGCAGCAGACATTTACATTTGAATGATGCTCAACCCAGTGGAGAA 960
Db 1022 GACTGATGATGATGCGCAGCAGACATTTATTTGAAATGATGCTCAACCCAGTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCTAGAACCAAAACAGCA 1141
QY 1021 TTTGACTTAATTTTAAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTGAATTTTAAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1201
QY 1081 ACAAGAACCTGCAAGAAATTTTGTCTTTTCAATTTGGCTGGAACTATGTTAAAGTTTGCATGCT 1140
Db 1202 ACAAGAACCTGCAAGAAATTTTGTCTTTTCAATTTGGCTGGAACTATGTTAAAGTTTGCATGCT 1261
QY 1141 ATTTGCTGTGTTTAAACAGCAAGGAGCAAAAGAAATACCTGTTTATTTTATGTGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGCAAGGAGCAAAAGAAATACCTGTTTATTTTATGTGACCA 1321
QY 1201 GCCAGACAGGCAAAATCTATTATTGCAAGCAAGCAATAGCAGCAAGCTGTGGCAATTTGGT 1260
Db 1322 GCCAGTACAGGCAAAATCTATTATTGCAAGCAAGCAATAGCAGCAAGCTGTGGCAATTTGGT 1381
QY 1261 TGTATAATGCGAGCAATGTAACTTTTCAATTTTATGATGTATCAACCAAGAACTTGAT 1320
```

Db 1382 TGTATATGACGAAATGGAATCTTCCATTCAATGACTGTACCAACAAGAACTTAATT 1441  
Qy 1321 TGGGTAGAAGAGCTGTAATCTTTGGACAGCAAGTAAACAGATTTAAAGCCATTTGCTCT 1380  
Db 1442 TGGGTGAAGAGCTGTAATCTTTGGACAAAGTAAACAGATTTAAAGCCATTTGCTCT 1501  
Qy 1381 GGTCAAGCTATTCGATGATCAAAAGGAAGGAGGAGCAACAGATTTGAACCAACCA 1440  
Db 1502 GGTCAAAACAAATTCGATGATCAAAAGGAAGGAGGAGCAACAGATTTGAACCAACCA 1561  
Qy 1441 GTCATATGACCAAAATGAGAACATTTACAGTGTGTGAGATAGGCTCGGAAGAAAGACCA 1500  
Db 1562 GTCATATGACCAAAATGAGAACATTTACAGTGTGTGAGATAGGCTCGGAAGAAAGACCA 1621  
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAACACATCTTGCCT 1560  
Db 1622 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAACACATCTTGCCT 1681  
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGTTAAG 1620  
Db 1682 GGTGACTTTGGTTGGTTGACAAACATGANTGGCCCATGATTTGTGCTTGGTTGTTAAG 1741  
Qy 1621 AATGGTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG 1680  
Db 1742 AATGGTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG 1801  
Qy 1681 TCAGAAACTGGGCGGAGCAAAAGGTGCCAACTCTATAAATTTACTAGTTCGGCACGC 1740  
Db 1802 ACGGAAACTGGGCGGAGCGAAGGTGCCGACTCTCTATAAATTCACCTAGTTCGGCACGC 1861  
Qy 1741 TCACCAATTCAGACACCGAAAGATGACGCTCTAGCCAGCAACATATGCACTAACTCCACTT 1800  
Db 1862 TCACCAATTCAGACACCGAAAGATGACGCTCTAGCCAGCAACATATGCACTAACTCCACTT 1921  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTGTTGGC 1860  
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTGTTGGC 1981  
Qy 1861 GGCATCTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920  
Db 1982 GGCATCTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 2041  
Qy 1921 CTGAGCCCACTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCGGTGGGGAACCG 1980  
Db 2042 CTGAGCCCACTGGTCAGAGATCGAGGAGGATTTGAAAGCGTCTCGGTGGGGAACCG 2101  
Qy 1981 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGACTAA 2019  
Db 2102 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGACTAA 2140

RESULT 13  
HOU34255  
LOCUS  
DEFINITION Hamster parvovirus DNA.  
ACCESSION U34255  
VERSION U34255.1 GI:1464792  
KEYWORDS  
SOURCE Hamster parvovirus  
ORGANISM Hamster parvovirus  
REFERENCE 1 (bases 1 to 4773)  
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,  
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.  
TITLE Molecular characterization of newly recognized rodent parvoviruses  
J. Gen. Virol. 77 (Pt 5), 895-911 (1996)  
96201434  
MEDLINE  
PUBMED 8609486  
REFERENCE 2 (bases 1 to 4773)  
Besselsen,D.G.  
AUTHORS  
Direct Submission  
TITLE Submitted (17-AUG-1995) David G. Besselsen, Department of  
Veterinary Pathology, University of Missouri-Columbia, W213

FEATURES  
source Veterinary Medicine Building, Columbia, MO 65211, USA  
Location/Qualifiers  
1. .4773  
/organism="Hamster parvovirus"  
/mol\_type="genomic DNA"  
/specific\_host="Mesocricetus auratus"  
/db\_xref="taxon:42843"

ORIGIN

Query Match 92.2%; Score 1862.2; DB 14; Length 4773;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
Qy 1 ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGAGAACCAACCTGGTTAAAGGAAAAA 60  
Db 122 ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGAGAACCAACCTGGTTAAAGGAAAA 181  
Qy 61 AGTAAACCAAGAGCTGCTCATTTGTTTTAAAAATGAAATGTTCACTGAATGGAAAA 120  
Db 182 AGTAAACCAAGAGCTGCTCATTTGTTTTAAAAATGAAATGTTCACTGAATGGAAAA 241  
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGCTGGAATCTTTACAA 180  
Db 242 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGCTGGAATCTTTACAA 301  
Qy 181 CGAGGAGCGGAAACTACTTTGGGACCAAGAGCGAGACATGGAATGGGAAAAACACAGTGGAT 240  
Db 302 CGAGGAGCGGAAACTACTTTGGGACCAAGAGCGAGACATGGAATGGGAAATCTTCAGTGGAT 361  
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTGATTCCTTTGTTTAAATAATGTTTATTTGAA 300  
Db 362 GAACTAAACCAAAAAAGCAAGTATTCATTTGATTCCTTTGTTTAAATAATGTTTATTTGAA 421  
Qy 301 GTCTTTAAACAAAGAAATATATTTCTGCTGGTATGTTTAAATGTTTGGCAACATGAATGG 360  
Db 422 GTCTCGATCAAAAGAACATAGCACCTAGTATGTTACTTTGTTTGTACGATGAATGG 481  
Qy 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420  
Db 482 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 541  
Qy 421 CAAGGGAATGTTGGAGAGGCAACTAATTTGTTTACTTGGAGCAGATGTTGTTTAAACAGCC 480  
Db 542 CAAGGGAATGTTGGAGAGGCAACTAATTTGTTTACTTGGAGCAGATGTTGTTTAAACAGCC 601  
Qy 481 TGTAAATGTGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAAT 540  
Db 602 TGTAGTGTGAGCTATTAACAGCTGAAAGAAATTAAGCTGAGAGAGATAGCGGAAGACAA 661  
Qy 541 GAGTGGTTACTCTACTTATTAAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGT 600  
Db 662 GAATGGTCACTTTGCTTACTTATAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGT 721  
Qy 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTTAAAGAAAAATAGCAGTGT 660  
Db 722 GTTGTGTTGGAAATATAGTTGCTTACTTCTTTTATCCAAAGAAAGAAATATGACCAGT 781  
Qy 661 CCACCAAGAGACGAGCGTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720  
Db 782 CCACCAAGAGACGAGCGTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 841  
Qy 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGTGATGACATCGCGCCAGAAACG 780  
Db 842 AAAGAGGCGGAAACGCCATCTAGTGAGCAAACTATACACTGTGATGACATCGCGCCAGAAACG 901  
Qy 781 GTTGAACCAACAGTAAACCACTGGCGAGGAAACTAAGCGGCGAGAAATTCAACTTAAAAA 840  
Db 902 GTTGAACCAACAGTAAACCACTGGCGAGGAAACTAAGCGGCGAGAAATTCAACTTAAAAA 961  
Qy 841 GAAGTTTCTATTAAAAACTACACTTAAAGAGCTCGTGCATAAAAAGAGTAACTCACCAGAG 900  
Db 962 GAGGTCTCTATTAAAAACCACTTAAAGAGCTCGTGCATAAAAAGAGTAACTCACCAGAA 1021

Qy	901	GACTGGATGATGATGCAGCCAGACAGTTACATTGAAATGATGGCTCAACAGGTGGAGAA	960
Db	1022	GACTGGATGATGATGCAGCCAGACAGTTACATTGAAATGATGGCTCAACAGGTGGAGAA	1081
Qy	961	AACCTGCTGAAAAATACGCTTAGAGATTGTGACACTTAACCTCTAGCGAGAACCAAAAACAGCA	1020
Db	1082	AACCTGCTGAAAAATACGCTTAGAGATTGTGACACTTAACCTCTAGCGAGAACCAAAAACAGCA	1141
Qy	1021	TTTTGACTTAATTTTAGAAAAAGCTGAAACAGACAGAACTAACCAACTTTTCACCTGCCTGAC	1080
Db	1142	TTTTGACTTAATTTTAGAAAAAGCTGAAACAGACAGAACTAACCAACTTTTCACCTGCCTGAC	1201
Qy	1081	ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTTCGCCATGCT	1140
Db	1202	ACHAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTTCGCCATGCT	1261
Qy	1141	ATTGTGCTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACCTGTTTATTTTCATGACCA	1200
Db	1262	ATTGTGCTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACCTGTTTATTTTCATGACCA	1321
Qy	1201	GCCAGCACAGGCAAACTATTTATGACAGAGCCATAGCACAGAGTTGCGAATGTTGGT	1260
Db	1322	GCCAGTACAGGCAAACTCCATTTATGCAAGAGCCATAGCACAGGAGTTGTTAATGTTGGT	1381
Qy	1261	TGCTATTAATGCAGCAATGTAAACTTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT	1320
Db	1382	TGCTATTAATGCAGCAATGTGAACTTTCCATTTTAAATGACTGTACCAACAACAACCTGATT	1441
Qy	1321	TGGGTAGAGAAGCTGTGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTGTGCTCT	1380
Db	1442	TGGGTAGAGAAGCTGTGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTGTGCTCT	1501
Qy	1381	GGTCAAGCTATTCGATTTGATCAAAAAGGAAAGGCGAGCAACACAGATTGAACCAACACCA	1440
Db	1502	GGTCAAACTATTCGCAATTTGATCAAAAAGGAAAGGCGAGCAACACAGATTGAACCAACACCA	1561
Qy	1441	GTCAATCATGACCACAAAATGAAACATTTACAGTGTGAGAACTAGGCTCGGAAGAAAGACCA	1500
Db	1562	GTCAATCATGACCACAAAATGAAACATTTACAGTGTGAGAACTAGGCTCGGAAGAAAGACCA	1621
Qy	1501	GAAACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCAATCTAAACACATACCTTGCTCT	1560
Db	1622	GAAACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCAATCTTAAACACATACATTCGCT	1681
Qy	1561	GGTGACTTTGGTTGTGACAAAATGCAATGGCCCATGATTGTGCTTGGTTGGTAAAG	1620
Db	1682	GGTGACTTTGGTTGTGACAAAATGCAATGGCCCATGATTGTGCTTGGTTGGTAAAG	1741
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTTAAATGGGGCAAGTTTCTGATTGG	1680
Db	1742	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTTAAATGGGGCAAGTTTCTGATTGG	1801
Qy	1681	TCAGAAAACTGGGCGGAGCCAAAGGTGCAACTCCTATTAATTTACTAGGTTTGGGACGC	1740
Db	1802	ACGAAAACTGGGCGGAGCCAAAGGTGCGGACTCCTATAATTCACCTAGGTTTGGGACGC	1861
Qy	1741	TCACCATTCAGCACACCGAAAGTACCGCTCTCAGCGAGACTATGACATAACTTCCACTT	1800
Db	1862	TCACCATTCAGCACACCGAAAGTACCGCTCTCAGCGAGACTATGACATAACTTCCACTT	1921
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAAACTACTCCTGTTGCG	1860
Db	1922	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAAACTACTCCTGTTGCG	1981
Qy	1861	GGCACTGCAGAAAACCCAGAAACATCTGGGGAAAGCTGGTTCCAAAGCTGCGAAGATGGTCAA	1920
Db	1982	GGCACTGCAGAAAACCCAGAAACATCTGGGGAAAGCTGGTTCCAAAGCTGCGAAGATGGTCAA	2041
Qy	1921	CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTGAGAGCGTGTCTCGGTGCGGAAACCG	1980
Db	2042	CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTGAGAGCGTGTCTCGGTGCGGAAACCG	2101
Qy	1981	TTGAAGAAAGACTTCAGGCGAGCCGTGAACTTGGCACTAA	2019

```

Db          2102 TTGAAGAGAGACTTCACGCGAGCCGCTGAACCTGGACTAA   2140
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 14
MVU34256    MVU34256             4761 bp      DNA       linear     VRL 21-AUG-1996
LOCUS      Mice minute virus DNA.
DEFINITION U34256
ACCESSION  U34256.1 GI:1464795
VERSION    .
KEYWORDS
SOURCE     Mice minute virus
ORGANISM   Mice minute virus
REFERENCE  1 (bases 1 to 4761)
AUTHORS   Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr., and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE  2 (bases 1 to 4761)
AUTHORS   Besselsen,D.G.
Direct Submission
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
JOURNAL
FEATURES
source     1..4761
            /organism="Mice minute virus"
            /mol_type="genomic DNA"
            /strain="Cutter"
            /specific_host="Mesocricetus auratus"
            /db_xref="taxon:10794"
ORIGIN
Query Match           91.9%; Score 1855.8; DB 14; Length 4761;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY  1 ATGCCTCGAATGCTTACTCTGTATGAAGTTTGGGAGCAACAACCTGGTTAAAGAAAAA 60
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  122 ATGCCTCGAATGCTTACTCTGTATGAAGTTTGGGAGCAACAACCTGGTTAAAGAAAAA 181
QY  61 AGTAACGAGGAAGTGTTCTCATTTTTGTTTTTAAAATAAGAAAATGTTCAACTGAAATGGAAAA 120
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  182 AGTAACGATTAGTATTCTCATTTGTTTTTAAAATAAGAGATGTTCAATTGAATGGAAAA 241
QY  121 GATATCGGATGAATAGTTTACAAAAAGAGCTGCAGAGGACGAGCTGAATCTTTACAA 180
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  242 GATATCGATGAATAGTTTACAGAAAGGAGCTGCAAGAGGACGAGCTAAAATCTTTACAA 301
QY  181 CGAGAGCGGAACCTACTCTGGGCCAACGAGGACATGGAATCGGAAACCACAGTGGAT 240
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  302 CGAGAGCGGAACCTACTCTGGGCCAACGAGGACATGGAATCGGAAATCTTTAGTGGAT 361
QY  241 GAAATGACCAAAAGCAAGTATTCATTTTTGTTCTTTTGGTTAAAAAATGTTTATTTGAA 300
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  362 GAACTAACCAAGAAGCAAGTATTCATTTTTGACTCTTTTAGTTAAAAAGTGTTTATTTGAA 421
QY  301 GTGCTTTACAAAGAAATATATTTTCCTGGTGCAATGTTAAITGGTTTGTGCACATGAATGG 360
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  422 GTGCTTAAGTACAAAGAACATAGCTCCTAGTGTATTAATTTGGTATGTGCAGCATGAATGG 481
QY  361 GGAAAGACCAAGGCTGSCACTGCCATGCTACTAAATTGGAGAAAAGGACTTTAGTCAAGCT 420
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  482 GGAAAGACCAAGGCTGSCATTTGCCATTGCAATGCTACTAATTGGAGCAAGACTTTAGCCAAGCT 541
QY  421 CAAGGGAATGGTGAGCAAGCAACTAAATGTTTTTCTGGAGCAGATGGTTGGTAAACGCC 480
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  542 CAAGGAAGTGGTGAGGAAGGCAGCTAAATGTTTTTACTTGGACGAGATGGTTGGTAAACGCC 601
QY  481 TGTAATGTGGAACCTAACACCAAGCTGAAGAANAATTAAACTAAGAGAAATAGCAGACAAAT 540

```

```
Db 602 TGAGTGTGAGCTATCACCAGCGGAAAGAAATTAAGCTGAGAGAAATAGCGGAAGACCAA 661
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGAGACTATACCAAGTGT 600
Db 662 GAATGGGTCACTTGTCTTACTTATAAGCATAAGCAAAACCAAAAGAGACTATACCAAGTGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCAGTAGT 660
Db 722 GTTGTCTTTGGAATATGATGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCAGTAGT 781
Qy 661 CCACCAAGAGACGAGGCTATTTTCTPAGCAGTGAAGTCTCTGCTGGAAAGCTAACTTTTAA 720
Db 782 CCACCAAGGAGCGAGGCTATTTTCTPAGCAGTGAAGTCTCTGCTGGAAAGCTAACTTTTAA 841
Qy 721 AAAGAAAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 842 AAAGAAAGCGGAAACGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 901
Qy 781 GTTGAACACACAGTAACCACTGGCGAGGAAACTTAAGCGCGGAGAAATTTCAAACTAAAGAA 840
Db 902 GTTGAACACACAGTAACCACTGGCGAGGAAACTTAAGCGCGGAGAAATTTCAAACTAAAGAA 961
Qy 841 GAAGTTTCTATTAAACTACACTTAAGAGCTGTGTCATAAAGAGTAACCTCACCAGAG 900
Db 962 GAGGTTTCTATTAAACCACTTAAGAGCTGTGTCATAAAGAGTAACCTCACCAGAA 1021
Qy 901 GACTGGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACAGAGTGGAGAA 960
Db 1022 GACTGGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAGCCAGGTGGAGAA 1081
Qy 961 AACTGCTGAAAAATACGCTAGAGATTTGTGACATACTAAGCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACTGCTGAAAAATACGCTAGAGATCTGTACGCTAAGCTAGCTAGAGAACCAAAACAGCC 1141
Qy 1021 TTTGACTTAATTTAGAAAGCTGAACAGGAGGCAAAAGAAATCTGTTTTATTTATGAGACCA 1080
Db 1142 TTTGACTTAATTTAGAAAGCTGAACAGGAGGCAAAAGAAATCTGTTTTATTTATGAGACCA 1201
Qy 1081 ACAAGAACTGCAAGATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGCCTATGCT 1140
Db 1202 ACAAGAACTGTAAGATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGCCTATGCT 1261
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTATTTATGAGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTATTTATGAGACCA 1321
Qy 1201 GCCAGCAGGCAAACTCTATTATGCAAGCCATAGCAAGCAGTGGCAATGTTGCT 1260
Db 1322 GCCAGTACAGGCAAACTCTATTATGCAAGCCATAGCAAGCAGTGGCAATGTTGCT 1381
Qy 1261 TGCTATAATGAGCAAACTGTAATTTTCCATTTTAACTGATGACCTGACCAAGAACTTGATT 1320
Db 1382 TGCTATAATGAGCAAACTGTAATTTTCCATTTTAACTGATGACCTGACCAAGAACTTGATT 1441
Qy 1321 TGGGTAGAAAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1442 TGGGTAGAAAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1501
Qy 1381 GGTCAAGCTATTGCAATTTGATCAAAAGGAAAGGCGAGCAAAAGATTTGAACCAACCA 1440
Db 1502 GGTCAAACTATTGCAATTTGATCAAAAGGAAAGGCGAGCAAAAGATTTGAACCAACCA 1561
Qy 1441 GTCATATGACCAAAATGAAACATTTACAGTGTGTGAGAAATGAGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATATGACCAAAATGAAACATTTACAGTGTGTGAGAAATGAGCTGCGAAGAGAGACCA 1621
Qy 1501 GAAACACACTCAACCAATCAGAGACAGAAATGCTTAACATTTCACTAACACATACCTTGCTCT 1560
Db 1622 GAAACACACTCAACCAATCAGAGACAGAAATGCTTAACATTTCACTAACACATACCTTGCTCT 1681
Qy 1561 GGTGACTTTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTGCTGGTGAAG 1620
```

```
Db 1682 GGTGACTTTTGGTTGTTGACAAAAATGAATGCCCATGATTTGTGCTTGTTGTTAAAG 1741
Qy 1621 AATGGTATCAATCTACCTACCTAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1742 AATGGTATCAATCTACCTACCTAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1801
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGCTTCGGCAACG 1740
Db 1802 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGCTTCGGCAACG 1861
Qy 1741 TCACCAATTCACGACACCGAAAAAGTAGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 1862 TCACCAATTCAGACACCGAAAAAGTAGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1921
Qy 1801 GCATCGGATCTCGAGACCTGCTTTAGAGCCTTTGAGAGCACAACAAATCTCTCTGTTGCG 1860
Db 1922 GCATCGGATCTCGAGACCTGCTTTAGAGCCTTTGAGAGCACAACAAATCTCTCTGTTGCG 1981
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGCGGAGAGCTGTTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1982 GGCACCTGCAGAAACCCAGAACACTGCGGAGAGCTGTTTCCAAAGCCTGCCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCTCGGTGGGAAACCG 1980
Db 2042 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCTCGGTGGGAAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGGACTAA 2140

RESULT 15
LOCUS CQ786765 5121 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 1 from Patent WO2004018689.
ACCESSION CQ786765
VERSION CQ786765.1 GI:45721778
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Iggo,R. and Mallerba,M.
TITLE Anti-neoplastic viruses.
JOURNAL Patent: WO 2004018689-A 1 04-MAR-2004;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
source 1: 5121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Parvovirus H1\WMV (pH1) - Parvovirus H1 with
promoter P4 and left hairpin from WMV"

ORIGIN
Query Match 87.7%; Score 1771; DB 6; Length 5121;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
Qy 1 ATCGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTTAAAGGAAAAA 60
Db 282 ATCGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTTAAAGGAAAAA 341
Qy 61 AGTAAACAGGAAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGGAAAA 120
Db 342 AGTAAACAGGAAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGGAAAA 401
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGAGGACGAGCTGGAATCTTTTACAA 180
Db 402 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGAGGACGAGCTGGAATCTTTTACAA 461
Qy 181 CGAGGAGCGAAACTACTTGGGACCAAGCAGGACATGGAATGGGAACCAACAGTGGAT 240
```

Db	462	CGAGGACGGAAACTACTTTGGGACCAAGCGAGGACATGGTAATGGGAACACACAGTGGAT	521
Qy	241	GAATGACCAAAAAGCAAGTATTCATTTTTTGATCTCTTTGGTTAAAAAATGTTTATTTGAA	300
Db	522	GAATGACCAAAAAGCAAGTATTCATTTTTTGATCTCTTTGGTTAAAAAATGTTTATTTGAA	581
Qy	301	GTGCTTAAACAAAGAATATATTTCTGTGTATGTTAATTTGGTTTGTGCAACATGAATGG	360
Db	582	GTGCTTAAACAAAGAATATATTTCTGTGTATGTTAATTTGGTTTGTGCAACATGAATGG	641
Qy	361	GGAAAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAAGCACTTTAGTCAAGCT	420
Db	642	GGAAAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAAGCACTTTAGTCAAGCT	701
Qy	421	CAAGGGAAATGGTGGAGAAAGGCAACTAAATGTTTTACTTGGAGCAGATGTTGGTTAACAGCC	480
Db	702	CAAGGGAAATGGTGGAGAAAGGCAACTAAATGTTTTACTTGGAGCAGATGTTGGTTAACAGCC	761
Qy	481	TGTAATGTGCAACTACACACAGCTGAAGAAATTTAAACTAAAGAGAAATAGCAGAAGCAAT	540
Db	762	TGTAATGTGCAACTACACACAGCTGAAGAAATTTAACTTAAGAGAAATAGCAGAAGCAAT	821
Qy	541	GAGTGGGTTTACTCTTACTTATTATAGCAATAAGCAAAACCAAAAAAGACTATACCAAGTGT	600
Db	822	GAGTGGGTTTACTCTTACTTATTATAGCAATAAGCAAAACCAAAAAAGACTATACCAAGTGT	881
Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTTTAACTTAAAGAGAAATAGCACTAGT	660
Db	882	GTTCCTTTTGGAAACATGATGCTTACTATTTTTTAACTTAAAGAGAAATAGCACTAGT	941
Qy	661	CCACCAAGAGACGGAGGCTATTTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTTTA	720
Db	942	CCACCAAGAGACGGAGGCTATTTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTTTA	1001
Qy	721	AAAGAGCGGAGCGCCATCTTAGTGAGCAAACTATACCTGATCACTACGCGGCAGAAAGC	780
Db	1002	AAAGAGCGGAGCGCCATCTAGTGAGCAAACTGTATCTGATGAGATGAACACGAAGACG	1061
Qy	781	GTTCGAACACACAGTAAACCACTGCGCGAGAAACTAAGCGCGCGAGAAATTCAAACTAAAAA	840
Db	1062	GTGAGACCAACAGTGAACCACTGACACAGGAGCTTAAGCGCGCGAGAAATTCAAACTAGAG	1121
Qy	841	GAAGTTTCTATTAATAAATCTACCTTAAAGAGCTGGTGATTAATAAGCTTAACCTCACAGAG	900
Db	1122	GAGGTCTCGATTAATAAACCACTCAAAAGAGTTGGTACATATAAAGAGTAACTCACAGAA	1181
Qy	901	GACTGGATGATGATGCGAGCAGACAGTTTACATTTGAAATGATGCTCAACAGGTGGAGAA	960
Db	1182	GACTGGATGATGATGCGAGCAGACAGTTTACATTTGAAATGATGCTCAACAGGTGGAGAA	1241
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACATACTCTAGCCAGAACCAAAACAGCA	1020
Db	1242	AACCTGCTTAAAAATACACTAGAGATCTGTACACTGACTCTAGCAAGAACCAAAACAGCC	1301
Qy	1021	TTTGACTTAAATTTTAAAAAGCTGAACACAGCAAACTTAACCAACTTTTCACTGCCTGAC	1080
Db	1302	TTTGACTTAAATTTTAAAAAGCTGAACACAGCAAACTTAGCCAACTTTTCCATGGCTAGC	1361
Qy	1081	ACAAGAACTGCGAGAAATTTTTGCTTTTTCATGGCTGGAACTATGTTAAAGTTTGCCTAGCT	1140
Db	1362	ACCAGAACCTGTAGAAATCTTTGCTGAGCATGGCTGGAACTATATTAAGTCTGCCATGCC	1421
Qy	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCATGACCA	1200
Db	1422	ATCTGTGTGTGCTGAAATAGACAGGAGGCAAAAGAAACCTGTGCTCTTTTCAGGACCA	1481
Qy	1201	GCCAGCACAGGCAAAATCTATTATTGCAACAGCCATAGCAACAGCAGTTGGCAATTTGGT	1260
Db	1482	GCAGCACAGGCAAAATCTATTATTGCAACAGCCATAGCAACAGCAGTTGGTAAATTTGGT	1541
Qy	1261	TGCTATAATGCGCCAAATGTAACCTTTCCATTTAATGATGTACCAACAGAACTTCATT	1320
Db	1542	TGTTACAATGCTGCCAATGTGAACCTTCATTTAATGATGTACCAACAAACCTTGAT	1601

Qy	1321	TGGGTAGAGAACTGGTAACTTTGGACAGCAAGTAAACCACTTTAAAGCCATTTTGCTCT	1380
Db	1602	TGGGTGGAAGAAGCTGGTAACTTTGGCCACAGCAAGTAAACCAATTTAAAGCTATTTGTCT	1661
Qy	1381	GGTCAAGCTATTCCGCTATTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGAACCAACACCA	1440
Db	1662	GGCCAAACCAATAGCATTTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGAACCAACACCA	1721
Qy	1441	GTCATCATGACCAAAATGAGAACTTACATGTGCTCAGAAATAGGCTGCCAAGAGAAAGACCA	1500
Db	1722	GTTATTATGACCACCAACGAGAACTTACCGTGGTTAGAAATAGGCTGTGAGGAAAGACCA	1781
Qy	1501	GAACACACTCAACCAATCACAGACAGAACTGCTTAACTCATCTAACACATACCTTTGCCCT	1560
Db	1782	GAACACACTCAACCAATCACAGACAGAACTGCTTAACTCATCTAACACATACCTTTGCCCT	1841
Qy	1561	GGTGACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGCTTGGTTGGTAAAG	1620
Db	1842	GGTGACTTTGGTTTGGTTGGATAAGACAGCAATGGCCCTCTGATCTGTGCTTGGTTGGTGAAG	1901
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCCTGATTTGG	1680
Db	1902	AATGGTTACCAATCTACCATGGCTTGTTCCTGTGCTTAAATGGGGCAAAAGTTCCTGATTTGG	1961
Qy	1681	TCAGAAAACTGGGGCGGAGCCAAAGGTGCCAACTCCTATAAAATTTAATGAGTTTCGGCAAGC	1740
Db	1962	TCAGAGGACTGGGGCGGAGCCGAGCTAGACACTCTATAAATTCGGTAGGTTTCAATGGCGC	2021
Qy	1741	TCACCATTCACGACACACGAAAGTACGCCCTTCAGCCAGACACTATGCACATACTCCACATT	1800
Db	2022	TCACCATCTCTGACTCCGAGAAAGTACGCCCTTCAGCCAAACACTACGCTCTTATCTCCACTT	2081
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACAAATATCTCGTTTGGC	1860
Db	2082	GCACTCGGACCTTGGGACCTTAGAGCTCTAGAGCCTTTGGAGCACACAAATATCTCGTTTGGC	2141
Qy	1861	GGCACTCGAGAAACCCAGAACACTTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA	1920
Db	2142	GGCACTCGAGCAAGCCAAAACACTTGGGGAGGCTGGTTCCACAGCCTGCCAAGGTGCTCAA	2201
Qy	1921	CTGAGCCCAACTTGTGTACAGATACGAGGAGGATTTGAGAGCGTCTCTCGGTGGGAAACCG	1980
Db	2202	CGAGGCCCAACTGTGTCCAGATCGAGGCGGATTTGAGAGCTTCTCTCAGTCAAGAACAG	2261
Qy	1981	TTTGAAGAAGACTTCAGCGAGCGCGCTGAACTTTGAGCTAA	2019
Db	2262	TTTGGAGAGCGACTTCAACGAGGAGCTGACCTTTGAGCTAA	2300

Search completed: January 22, 2005, 21:19:34  
Job time : 8548.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds

(without alignments)  
11150.532 Million cell updates/sec

Title: US-10-069-056-16

Perfect score: 2019

Sequence: 1 atggcgggaatgcttactc.....agccgctgaacttggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*

1: Geneseqn1980s:\*

2: Geneseqn1920s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2019	100.0	2019	5	AAD02805 Parvoviru
2	2017.4	99.9	2019	5	AAD02797 Parvoviru
3	2017.4	99.9	5149	12	Adg39767 Minute vi
4	2015.8	99.8	2019	5	AAD02801 Parvoviru
5	2015.8	99.8	2019	5	AAD02803 Parvoviru
6	2015.8	99.8	2019	5	AAD02799 Parvoviru
7	1040.4	51.5	5049	2	AAT15311 Non-atten
8	1040.4	51.5	5049	2	AAT15312 Attenuate
9	1040.4	51.5	5049	2	AAT88324 Attenuate
10	1038.8	51.5	5049	2	AAT88321 Attenuate
11	1038.8	51.5	5049	2	AAT88320 Canine pa
12	901	44.6	3524	1	AAN40252 Sequence
13	307.4	15.2	374	6	ABQ95626 Tumour su
14	285.8	14.2	421	6	ABQ94779 Tumour su
15	284.8	14.1	423	6	ABQ94724 Tumour su
16	284.8	14.1	464	6	ABQ94719 Tumour su
17	284.8	14.1	473	6	ABQ94716 Tumour su
18	284.8	14.1	486	6	ABQ94707 Tumour su
19	284.8	14.1	491	6	ABQ94777 Tumour su
20	284.2	14.1	420	6	ABQ94787 Tumour su
21	283.8	14.1	343	6	ABQ94737 Tumour su

ALIGNMENTS

RESULT 1  
AAD02805  
ID AAD02805 standard; DNA; 2019 BP.  
XX  
AC AAD02805;

DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX

DE Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.  
OS Synthetic.

FH Key Location/Qualifiers  
FT CDS 1..2019  
FT /\*tag= a  
FT /\*product= "NS1 variant (T463A) protein"  
FT mutation replace(1387, A)  
FT /\*tag= b

XX EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

XX WPI; 2001-212717/22.

XX P-PSDB; AAY72710.

XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.

XX Claim 7; Page 27-30; 41pp; English.

Abq94739 Tumour su  
Abq94782 Tumour su  
Abq94718 Tumour su  
Abq94732 Tumour su  
Abq94765 Tumour su  
Abq94781 Tumour su  
Abq94740 Tumour su  
Abq95625 Tumour su  
Abq94778 Tumour su  
Abq94873 Tumour su  
Abq94867 Tumour su  
Abq94868 Tumour su  
Abq94874 Tumour su  
Abq94869 Tumour su  
Abq94880 Tumour su  
Abq94793 Tumour su  
Abq94807 Tumour su  
Abq94794 Tumour su  
Abq94800 Tumour su  
Abq94811 Tumour su  
Abq94791 Tumour su  
Abq94799 Tumour su  
Abq94796 Tumour su  
Abq94805 Tumour su



XX	The present sequence is a DNA encoding parvovirus non-structure protein 1 (NS1) variant (T463A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the cytotoxicity activity. These variants are useful as toxins for treating tumoural diseases. The variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-2003 to correct OS field.)									
XX	Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;									
QY	Query Match	100.0%;	Score 2019;	DB 5;	Length 2019;					
DB	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 2019;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGGACCAACCACTGGTTAAAGGAAAAA	60							
DB	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGGACCAACCACTGGTTAAAGGAAAAA	60							
QY	61	AGTAACCAAGGAAGTGTTCATTTGTTTAAAAATGAAATCTTCAACTGGAATGGAATA	120							
DB	61	AGTAACCAAGGAAGTGTTCATTTGTTTAAAAATGAAATCTTCAACTGGAATGGAATA	120							
QY	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGTAATCTTTACAA	180							
DB	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGTAATCTTTACAA	180							
QY	181	CGAGGAGCGAATCTACTTGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT	240							
DB	181	CGAGGAGCGAATCTACTTGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT	240							
QY	241	GAATGACCAAAAAAGCAAGTATTCATTTTGTATCTTTGTTAAAAATGTTTATTGAA	300							
DB	241	GAATGACCAAAAAAGCAAGTATTCATTTTGTATCTTTGTTAAAAATGTTTATTGAA	300							
QY	301	GTGCTTAAACCAAAAGAAATATTTCTCTGGTGATGTTAATTTGTTTGTGCAATGAATGG	360							
DB	301	GTGCTTAAACCAAAAGAAATATTTCTCTGGTGATGTTAATTTGTTTGTGCAATGAATGG	360							
QY	361	GGAAAGACCAAGCTGGCACTGCCATGACTAATTTGGAGGAAGGACTTTAGTCAAGCT	420							
DB	361	GGAAAGACCAAGCTGGCACTGCCATGACTAATTTGGAGGAAGGACTTTAGTCAAGCT	420							
QY	421	CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGTTAACAGCC	480							
DB	421	CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGTTAACAGCC	480							
QY	481	TGTAATGTGCAACTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540							
DB	481	TGTAATGTGCAACTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540							
QY	541	GAGTGGGTTACTCTTACTTATAGCATAAGCAACCAACCAAAAGACTATATCAAGTGT	600							
DB	541	GAGTGGGTTACTCTTACTTATAGCATAAGCAACCAACCAAAAGACTATATCAAGTGT	600							
QY	601	GTTCCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAAAATAGCACTAGT	660							
DB	601	GTTCCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAAAATAGCACTAGT	660							
QY	661	CCACCAAGAGACGGAGCTATTTTCTTAGCAGTGA CTCTGGCTGGAAAACTAATCTTTTAA	720							
DB	661	CCACCAAGAGACGGAGCTATTTTCTTAGCAGTGA CTCTGGCTGGAAAACTAATCTTTTAA	720							
QY	721	AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780							
DB	721	AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780							
QY	781	GTTGAAACCAAGTAACTGCGGAGGAACTAAGCGCGGAGAGTAATTTCAAACTAAAAA	840							
DB	781	GTTGAAACCAAGTAACTGCGGAGGAACTAAGCGCGGAGAGTAATTTCAAACTAAAAA	840							
QY	841	GAGGTTTCTATTAAAACTACATTAAGAGCTGGTGATTAAGAGTAATCTCACCGAG	900							
DB										

DB	841	GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGTCATAAAAAGAGTAACCTCACAGAG	900							
QY	901	GACTGATGATCATGAGCCAGACAGATTTACATTTGAAATGATGGCTCAACAGATGGAGAA	960							
DB	901	GACTGATGATCATGAGCCAGACAGATTTACATTTGAAATGATGGCTCAACAGATGGAGAA	960							
QY	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAGAACCAACAGCA	1020							
DB	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAGAACCAACAGCA	1020							
QY	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCACTAACTAACTTTTCACTGCCTGAC	1080							
DB	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCACTAACTAACTTTTCACTGCCTGAC	1080							
QY	1081	ACAAGAACCTGAGAGATTTTTCATGCTGGAATATGTTTAAAGTTTGGCCATGCT	1140							
DB	1081	ACAAGAACCTGAGAGATTTTTCATGCTGGAATATGTTTAAAGTTTGGCCATGCT	1140							
QY	1141	ATTGCTGCTGTTTAAACAGAGAGGCAAAAAGAAATACCTGTTTATTTTATGAGCA	1200							
DB	1141	ATTGCTGCTGTTTAAACAGAGAGGCAAAAAGAAATACCTGTTTATTTTATGAGCA	1200							
QY	1201	GCAGACAGGCAAACTCTATTATTCACAAAGCCATAGCACAGCTTGCATGTTTGT	1260							
DB	1201	GCAGACAGGCAAACTCTATTATTCACAAAGCCATAGCACAGCTTGCATGTTTGT	1260							
QY	1261	TGCTATAATGCAAGCAATGTAACCTTTTCACTTTAATGACTGTACCAACAGAACTTTGAT	1320							
DB	1261	TGCTATAATGCAAGCAATGTAACCTTTTCACTTTAATGACTGTACCAACAGAACTTTGAT	1320							
QY	1321	TGGGTAGAGAGAGCTGGTAACTTTTGACAGCAAGTAACAGTTTAAAGCATTTGCTCT	1380							
DB	1321	TGGGTAGAGAGAGCTGGTAACTTTTGACAGCAAGTAACAGTTTAAAGCATTTGCTCT	1380							
QY	1381	GGTCAAGCTATTTCGCATTCATCAAAAGGAAAGGAGGCAAGATTAACCAACCA	1440							
DB	1381	GGTCAAGCTATTTCGCATTCATCAAAAGGAAAGGAGGCAAGATTAACCAACCA	1440							
QY	1441	GTCTATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA	1500							
DB	1441	GTCTATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA	1500							
QY	1501	GAAACACTCAACCAATGAGAGACAGATGCTTAACTTATCATCTAACAATCTTGCCT	1560							
DB	1501	GAAACACTCAACCAATGAGAGACAGATGCTTAACTTATCATCTAACAATCTTGCCT	1560							
QY	1561	GGTGACTTTTGGTTTGGTTCACAAAAATGAATGSCCATGATTTGCTTGGTTGGTAAAG	1620							
DB	1561	GGTGACTTTTGGTTTGGTTCACAAAAATGAATGSCCATGATTTGCTTGGTTGGTAAAG	1620							
QY	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGTATGG	1680							
DB	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGTATGG	1680							
QY	1681	TCAGAAAACTGGGCGAGCAAAAGGTGCACTCTTATTAATTTACTAGTTTGGCAGCG	1740							
DB	1681	TCAGAAAACTGGGCGAGCAAAAGGTGCACTCTTATTAATTTACTAGTTTGGCAGCG	1740							
QY	1741	TCACCAATTCAGACACCGAAAGTAGCCCTCTCAGCCAGACTATGCACTAACTTCCACTT	1800							
DB	1741	TCACCAATTCAGACACCGAAAGTAGCCCTCTCAGCCAGACTATGCACTAACTTCCACTT	1800							
QY	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCACCAAAATCTCTCTGTGG	1860							
DB	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCACCAAAATCTCTCTGTGG	1860							
QY	1861	GGCACTGCGAAGAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTC	1920							
DB	1861	GGCACTGCGAAGAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTC	1920							
QY	1921	CTGAGCCCACTTGTTCAGAGATTCGAGGAGGATTTTCAGAGCGGTGCTTCGGTGGGAAACCG	1980							
DB	1921	CTGAGCCCACTTGTTCAGAGATTCGAGGAGGATTTTCAGAGCGGTGCTTCGGTGGGAAACCG	1980							



QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTTAA 2019  
DB 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTTAA 2019

RESULT 2  
AAD02797  
ID AAD02797 standard; DNA; 2019 BP.  
AC AAD02797;  
XX  
DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX  
DE Parvovirus non-structure protein 1 (NS1) wild-type DNA.  
XX  
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; ds.  
XX  
OS Parvovirus.

XX Key Location/Qualifiers  
XX CDS 1..2019  
FT /\*cag= a  
FT /product= "Parvovirus NS1 protein"

XX EPI077260-A1.  
XX  
XX 21-FEB-2001.  
XX  
XX 13-AUG-1999; 99EP-00115161.  
XX  
XX 13-AUG-1999; 99EP-00115161.  
XX  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;  
XX WPI: 2001-212717/22.  
XX P-PSDB; AAY72702.  
XX  
XX Novel parvovirus non-structure protein variant, useful for treating  
XX tumoral diseases, has a shifted equilibrium between DNA replication and  
XX transcription activities, and cytotoxic activity.  
XX  
XX Disclosure; Fig 1; 41pp; English.  
XX  
XX The present sequence is a wild type DNA encoding parvovirus non-  
XX structure protein 1 (NS1). The present invention relates to the variants  
XX of the parvovirus non-structure protein (NS1) having a shifted  
XX equilibrium between the DNA replication and transcription activities, and  
XX the cytotoxicity activity. These variants are useful as toxins for  
XX treating tumoural diseases. The variant DNAs are useful as vectors for  
XX gene therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
XX Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;

Query Match 99.9%; Score 2017.4; DB 5; Length 2019;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60  
DB 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60  
QY 61 AGTAACGAGAGTGTCTCATTGTTTTTAAATGAAATGTTCAACTGGAATGGAAAA 120  
DB 61 AGTAACGAGAGTGTCTCATTGTTTTTAAATGAAATGTTCAACTGGAATGGAAAA 120  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCGAGCTGGAATCTTTACAA 180  
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCGAGCTGGAATCTTTACAA 180

QY 181 CGAGGAGCGGAAACTACTTGGGACCAAAAGCGAGCAATGGAATGGGAAACACACAGTGGAT 240  
DB 181 CGAGGAGCGGAAACTACTTGGGACCAAAAGCGAGCAATGGAATGGGAAACACACAGTGGAT 240  
QY 241 GAAATGACCAAAAGCAAGTATTCTTTTGGTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAA 300  
DB 241 GAAATGACCAAAAGCAAGTATTCTTTTGGTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAA 300  
QY 301 GTGCTTAAACAAAGAAATATATTTCTCGTGTGATGTTAAATGTTTGGTAAATGTTTAAATGTTTAA 360  
DB 301 GTGCTTAAACAAAGAAATATATTTCTCGTGTGATGTTAAATGTTTGGTAAATGTTTAAATGTTTAA 360  
QY 361 GGAAGAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420  
DB 361 GGAAGAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420  
QY 421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTTAAACAGCC 480  
DB 421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTTAAACAGCC 480  
QY 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
DB 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
QY 541 GAGTGGTTACTCTTACTTATAGCATAAGCAAAACCAAAAGAAAGACTATACCAAGTGT 600  
DB 541 GAGTGGTTACTCTTACTTATAGCATAAGCAAAACCAAAAGAAAGACTATACCAAGTGT 600  
QY 601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAGAAAGAAATAGCACTAGT 660  
DB 601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAGAAAGAAATAGCACTAGT 660  
QY 661 CCACCAAGAGACGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAATTAACCTTTTAA 720  
DB 661 CCACCAAGAGACGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAATTAACCTTTTAA 720  
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780  
DB 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780  
QY 781 GTTGAACCACTAACCCTGCGAGGAACTAAGCGGCGAGATTAACCTAAGAAATTA 840  
DB 781 GTTGAACCACTAACCCTGCGAGGAACTAAGCGGCGAGATTAACCTAAGAAATTA 840  
QY 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 900  
DB 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 900  
QY 901 GACTGGATGATGATGAGCGAGACAGTTACATTTGAAATGATGCTCAACGAGTGGAGAA 960  
DB 901 GACTGGATGATGATGAGCGAGACAGTTACATTTGAAATGATGCTCAACGAGTGGAGAA 960  
QY 961 AACTGCTGAAAATAGCTAGAGATTTGTACACTAATCTAGCAGAAACCAAAACAGCA 1020  
DB 961 AACTGCTGAAAATAGCTAGAGATTTGTACACTAATCTAGCAGAAACCAAAACAGCA 1020  
QY 1021 TTTGACTTAAATTTAGAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080  
DB 1021 TTTGACTTAAATTTAGAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080  
QY 1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCGTGGAACTATGTTAAAGTTTGCATGCT 1140  
DB 1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCGTGGAACTATGTTAAAGTTTGCATGCT 1140  
QY 1141 ATTTGCTGTGTTTTAAACAGACAGGAGGCAAAAGAAATGTTTATTTTATGAGCA 1200  
DB 1141 ATTTGCTGTGTTTTAAACAGACAGGAGGCAAAAGAAATGTTTATTTTATGAGCA 1200  
QY 1201 GCCAGACAGCGCAAACTATTATTGCAAGCCATAGCAAGCAGTGGCAATGTTGGT 1260  
DB 1201 GCCAGACAGCGCAAACTATTATTGCAAGCCATAGCAAGCAGTGGCAATGTTGGT 1260

QY 1261 TGCTATAATGCAGCCAAATGTAATCTTCCATTAAATGACTGTACCAACAAAGAACTTGATT 1320  
Db 1261 TGCTATAATGCAGCCAAATGTAATCTTCCATTAAATGACTGTACCAACAAAGAACTTGATT 1320  
QY 1321 TGGGTAGAGAAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTCCTCT 1380  
Db 1321 TGGGTAGAGAAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTCCTCT 1380  
QY 1381 GGTCAAGCTATTGCGATTGATCAAAAGAGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440  
Db 1381 GGTCAAGCTATTGCGATTGATCAAAAGAGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440  
QY 1441 GTCATCATGACCAAAATAGAAATTCATGAGTGGTCAAGTATGAGTGGCGAAGAAAGCA 1500  
Db 1441 GTCATCATGACCAAAATAGAAATTCATGAGTGGTCAAGTATGAGTGGCGAAGAAAGCA 1500  
QY 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCAACACATCTTGCCCT 1560  
Db 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCAACACATCTTGCCCT 1560  
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCCCATGATTTGTGCTTGGTTGGTAAAG 1620  
Db 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCCCATGATTTGTGCTTGGTTGGTAAAG 1620  
QY 1621 AATGGTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCCTGATTGG 1680  
Db 1621 AATGGTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCCTGATTGG 1680  
QY 1681 TCAGAAACTGGGGGAGCAAGGTGCCAATCTTAAATTAATTAATTAATTAATTAATTAATTA 1740  
Db 1681 TCAGAAACTGGGGGAGCAAGGTGCCAATCTTAAATTAATTAATTAATTAATTAATTAATTA 1740  
QY 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACATT 1800  
Db 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACATT 1800  
QY 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGCTTGAGAGCACAAATATCTCTGTGCG 1860  
Db 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGCTTGAGAGCACAAATATCTCTGTGCG 1860  
QY 1861 GGCACCTGCAGAAACCCAGACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTCAT 1920  
Db 1861 GGCACCTGCAGAAACCCAGACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTCAT 1920  
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTGTTCGTTGCGGAACCG 1980  
Db 1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTGTTCGTTGCGGAACCG 1980  
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2019  
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2019

RESULT 3  
ADG39767  
ID ADG39767 standard; DNA; 5149 BP.  
XX AC ADG39767;  
XX AC ADG39767;  
XX AC ADG39767;  
DT 11-MAR-2004 (first entry)  
XX Minute virus from mouse genomic DNA sequence SEQ ID NO:29.  
DE parvovirus; rep; cap; DNA binding domain; capsid interacting domain;  
XX recombinant hybrid parvovirus particle;  
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;  
KW gene; ds.  
XX Mice minute virus.  
OS Mice minute virus.  
XX WO2003104392-A2.  
PN 18-DEC-2003.  
PD 18-DEC-2003.

XX 02-DEC-2002; 2002WO-US038423.  
PF 18-DEC-2001; 2001US-0341919P.  
XX (UINC-) UNIV NORTH CAROLINA.  
PA Samulski RJ, Rabinowitz JE;  
PI WPI; 2004-062324/06.  
XX New polynucleotides comprising parvovirus rep coding sequences and  
PT parvovirus cap coding sequences, useful in producing higher stocks of  
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a  
PT subject.  
XX Disclosure; SEQ ID NO 29; 115pp; English.  
XX The present invention describes a polynucleotide (I), comprising  
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The  
CC rep coding sequences encodes a DNA binding domain from a first  
CC parvovirus, and a capsid interacting domain from a parvovirus different  
CC from the first parvovirus. The cap coding sequence comprises sequences  
CC from the different parvovirus. Also described: (1) a vector comprising  
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and  
CC parvovirus cap coding sequences, where the rep coding sequences encode a  
CC DNA binding domain from a first parvovirus and a capsid interacting  
CC domain from a parvovirus different from the first parvovirus, the cap  
CC coding sequences comprise sequences from the different parvovirus, and  
CC the rep coding sequences are stably integrated into the genome of the  
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno  
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in  
CC gene therapy. The polynucleotide (I) can be used in producing higher  
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in  
CC the delivery of nucleic acids having biological effect to treat or  
CC ameliorate the symptoms associated with any disorder related to gene  
CC expression. The polynucleotide may be used to produce a parvovirus vector  
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.  
CC The parvovirus vector may also be used to provide an antisense nucleic  
CC acid to a cell in vitro or in vivo, or in diagnostic and screening  
CC methods. The present sequence is used in the exemplification of the  
CC present invention.  
XX Sequence 5149 BP; 1718 A; 1045 C; 1124 G; 1262 T; 0 U; 0 Other;  
SQ Query Match 99.9%; Score 2017.4; DB 12; Length 5149;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 261 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 320  
QY 61 AGTAACCGAGAGTGTCTCATTGTTTAAATAATGAAATGTTCACTCAATGGAATA 120  
Db 321 AGTAACCGAGAGTGTCTCATTGTTTAAATAATGAAATGTTCACTCAATGGAATA 380  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCAGAGGACGAGCTGAAATCTTTACAA 180  
Db 381 GATATCGGATGGAATAGTTTACAAAAAGAGCTCAGAGGACGAGCTGAAATCTTTACAA 440  
QY 181 CGAGGAGCGGAAACTACTCTGGGACCAAGAGGACGAGGACATGGAATGGGAAACCACTGGAT 240  
Db 441 CGAGGAGCGGAAACTACTCTGGGACCAAGAGGACGAGGACATGGAATGGGAAACCACTGGAT 500  
QY 241 GAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA 300  
Db 501 GAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA 560  
QY 301 GTGCTTAAACACAAAGAAATATTTCTGTTGATGTTTAAATGTTTGTGCAACATGAATGG 360  
Db 561 GTGCTTAAACACAAAGAAATATTTCTGTTGATGTTTAAATGTTTGTGCAACATGAATGG 620

QY 361 GGAAGACCAAGCGCTGGCACTGTCATTAATTCGAGGAAGGACTTTAGTCAAGCT 420  
Db 621 GGAAGACCAAGCGCTGGCACTGTCATTAATTCGAGGAAGGACTTTAGTCAAGCT 680  
QY 421 CAAGGGAATTTGGTGGAGAGGCACTAAATGCTTTTACTTGGAGCAGATGTTGGTAAACAGCC 480  
Db 681 CAAGGGAATTTGGTGGAGAGGCACTAAATGCTTTTACTTGGAGCAGATGTTGGTAAACAGCC 740  
QY 481 TGTAATGTCAACTTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
Db 741 TGTAATGTCAACTTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 800  
QY 541 GAGTGGTTACTTACTTACTTATTAAGCATTAAGCAACCAACCAAGAGACTATACCAAGTGT 600  
Db 801 GAGTGGTTACTTACTTACTTATTAAGCATTAAGCAACCAACCAAGAGACTATACCAAGTGT 860  
QY 601 GTTCTTTTGGAAACATGATTGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCACTAGT 660  
Db 861 GTTCTTTTGGAAACATGATTGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCACTAGT 920  
QY 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAGAACTAACTTTTAA 720  
Db 921 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAGAACTAACTTTTAA 980  
QY 721 AAAGAGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780  
Db 981 AAAGAGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 1040  
QY 781 GTTGAACCAACAGTAACCACTGCGCAGGAACTAAAGCGCGCAGAAATTCAAATTAAGAA 840  
Db 1041 GTTGAACCAACAGTAACCACTGCGCAGGAACTAAAGCGCGCAGAAATTCAAATTAAGAA 1100  
QY 841 GAAGTTTCTATTAAGAACTACACTTAAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG 900  
Db 1101 GAAGTTTCTATTAAGAACTACACTTAAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG 1160  
QY 901 GACTGGATGATGACGAGCAGACAGTTTACATTAAGAGTGGTCAACAGGTTGGAGAA 960  
Db 1161 GACTGGATGATGACGAGCAGACAGTTTACATTAAGAGTGGTCAACAGGTTGGAGAA 1220  
QY 961 AACTGCTGAAAATACGCTAGAGATTTGTACACTACTTAGCCAGAACCAAAACAGCA 1020  
Db 1221 AACTGCTGAAAATACGCTAGAGATTTGTACACTACTTAGCCAGAACCAAAACAGCA 1280  
QY 1021 TTTGACTTAATTTTAGAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080  
Db 1281 TTTGACTTAATTTTAGAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1340  
QY 1081 ACAAGAACCTGCGAAGATTTTGTCTTTTCACTGGGAACTATGTTAAAGTTTGGCATGCT 1140  
Db 1341 ACAAGAACCTGCGAAGATTTTGTCTTTTCACTGGGAACTATGTTAAAGTTTGGCATGCT 1400  
QY 1141 ATTTGCTGTTTAAACAGACAGGAGGCAAGAAATGCTGTTTATTTATTCATGGACCA 1200  
Db 1401 ATTTGCTGTTTAAACAGACAGGAGGCAAGAAATGCTGTTTATTTATTCATGGACCA 1460  
QY 1201 GCCAGCAGCAGCAATCTATTATTGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT 1260  
Db 1461 GCCAGCAGCAGCAATCTATTATTGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT 1520  
QY 1261 TGCTATAATGAGCCAAATGAACTTTTCCATTTTAACTGATGTAACCAAGAACTTGAT 1320  
Db 1521 TGCTATAATGAGCCAAATGAACTTTTCCATTTTAACTGATGTAACCAAGAACTTGAT 1580  
QY 1321 TGGGTAGAAGAGCTGTAACCTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380  
Db 1581 TGGGTAGAAGAGCTGTAACCTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1640  
QY 1381 GGTCAAGCTATTGCAATTTGATCAAAAAGGAAAGGAGCAAGAAACAGATTGAACCAACCA 1440  
Db 1641 GGTCAAGCTATTGCAATTTGATCAAAAAGGAAAGGAGCAAGAAACAGATTGAACCAACCA 1700  
QY 1441 GTCATCATGACCACAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGACCA 1500

Db 1701 GTCATCATGACCACAAATGAGAACTATTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGACCA 1760  
QY 1501 GAACACACTCAACCAATTCAGAGACAGAAATGCTTAAACATTCATCTTAAACACATACCTTGCT 1560  
Db 1761 GAACACACTCAACCAATTCAGAGACAGAAATGCTTAAACATTCATCTTAAACACATACCTTGCT 1820  
QY 1561 GGTGACTTTGGTTGGTTGTCAGAAAATGAATGAGCCCAAGTATTTGCTGCTTGGTGGTAAAG 1620  
Db 1821 GGTGACTTTGGTTGGTTGTCAGAAAATGAATGAGCCCAAGTATTTGCTGCTTGGTGGTAAAG 1880  
QY 1621 AATGGTTACCAATCTACCATGCGCAAGCTACTGCTGCTTAAATGGGGCAAGTTTCTGATGG 1680  
Db 1881 AATGGTTACCAATCTACCATGCGCAAGCTACTGCTGCTTAAATGGGGCAAGTTTCTGATGG 1940  
QY 1681 TCAGAAAACCTGGCGGAGCCAAAGGTGCGCAACTCTCTATAAATTTACTAGGTTTCGGCAGCG 1740  
Db 1941 TCAGAAAACCTGGCGGAGCCAAAGGTGCGCAACTCTCTATAAATTTACTAGGTTTCGGCAGCG 2000  
QY 1741 TCACCAATTCAGACACCGAAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 2001 TCACCAATTCAGACACCGAAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060  
QY 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACCTCTGTTGGC 1860  
Db 2061 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACCTCTGTTGGC 2120  
QY 1861 GGCACTGCAAGAACCCAGAACACTGGGAAAGCTGGTTTCCAAAGCTGCCAAGATGGTCAA 1920  
Db 2121 GGCACTGCAAGAACCCAGAACACTGGGAAAGCTGGTTTCCAAAGCTGCCAAGATGGTCAA 2180  
QY 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGCGAAACCG 1980  
Db 2181 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGCGAAACCG 2240  
QY 1981 TTGAAGAAAGACTTCAGGAGCGCGCTGAACTTGGACTAA 2019  
Db 2241 TTGAAGAAAGACTTCAGGAGCGCGCTGAACTTGGACTAA 2279

RESULT 4  
AAD02801 standard; DNA; 2019 BP.  
XX AAD02801;  
AC AC  
XX XX  
DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.  
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.  
XX Parvovirus.  
OS Synthetic.  
Key Location/Qualifiers  
CDS 1..2019  
FT /\*tag= a  
FT /product= "NS1 variant (T363A) protein"  
FT mutation replace(1187, A)  
FT /\*tag= b  
XX EP1077260-A1.  
PN 21-FEB-2001.  
XX 13-AUG-1999; 99EP-00115161.  
XX 13-AUG-1999; 99EP-00115161.  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

[illegible]



Db 1021 TTTGACTTAATTTTAAAAAGCTGAAACACGACAACTAAACCACTTTTCACTGCCTGAC 1080  
Qy 1081 ACAAGAACCTGCAGAAATTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1140  
Db 1081 ACAAGAACCTGCAGAAATTTTTCCTTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1140  
Qy 1141 ATTGCTGTGTTTAAACACAGAGGAGGAAAAAGAAATCTGTTTATTTCATGACCA 1200  
Db 1141 ATTGCTGTGTTTAAACACAGAGGAGGAAAAAGAAATCTGTTTATTTCATGACCA 1200  
Qy 1201 GCCAGCAGGCAATCTATATGCAACGCAATGACCAAGCAGTTGGCAATGTTGCT 1260  
Db 1201 GCCAGCAGGCAATCTATATGCAACGCAATGACCAAGCAGTTGGCAATGTTGCT 1260  
Qy 1261 TGCTATATGAGCCCAATGTAACCTTTCCATTTTAAAGCTGTACCAAGCAAGCACTTGATT 1320  
Db 1261 TGCTATATGAGCCCAATGTAACCTTTCCATTTTAAAGCTGTACCAAGCAAGCACTTGATT 1320  
Qy 1321 TGGGTAGAAGAGCTGTAATCTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTGCTCT 1380  
Db 1321 TGGGTAGAAGAGCTGTAATCTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTGCTCT 1380  
Qy 1381 GGTCAAGCTATTGCAATGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440  
Db 1381 GGTCAAGCTATTGCAATGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440  
Qy 1441 GTCATCATGACCAAAATGAGACATTACAGTGTGAGATAGCTGCGAAGAAAGACCA 1500  
Db 1441 GTCATCATGACCAAAATGAGACATTACAGTGTGAGATAGCTGCGAAGAAAGACCA 1500  
Qy 1501 GAACACACTCAACCAATCAGAGACAGAACTCTTAACATTCATTAACATACCTTGCCT 1560  
Db 1501 GAACACACTCAACCAATCAGAGACAGAACTCTTAACATTCATTAACATACCTTGCCT 1560  
Qy 1561 GGTGACTTTGGTTGTCAGAAATGAAAGGAGCAAGTATGCTGCTGTTGTTAAAG 1620  
Db 1561 GGTGACTTTGGTTGTCAGAAATGAAAGGAGCAAGTATGCTGCTGTTGTTAAAG 1620  
Qy 1621 AATGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG 1680  
Db 1621 AATGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG 1680  
Qy 1681 TCAGAAACTGGGCGAGCAAGGTCGCAACTCTTAAATTTTAAATTTTACGAGCTGGCAGCG 1740  
Db 1681 TCAGAAACTGGGCGAGCAAGGTCGCAACTCTTAAATTTTAAATTTTACGAGCTGGCAGCG 1740  
Qy 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGCACTATGCACTAATCCACTT 1800  
Db 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGCACTATGCACTAATCCACTT 1800  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG 1860  
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG 1860  
Qy 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGTTCAA 1920  
Db 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGTTCAA 1920  
Qy 1921 CTGAGCCCACTTGGTCAGAGATCGAGAGGATTTGAGCGCTGCTTCGTTGCGGACCG 1980  
Db 1921 CTGAGCCCACTTGGTCAGAGATCGAGAGGATTTGAGCGCTGCTTCGTTGCGGACCG 1980  
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 6

AAD02799

ID AAD02799 standard; DNA; 2019 BP.

XX

AC AAD02799;

XX

DT 06-AUG-2003 (revised)  
DT 31-MAY-2001 (first entry)  
XX Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.  
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;  
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.  
XX Parvovirus.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH CDS 1..2019  
FT /\*tag= a  
FT /product= "NS1 variant (S283A) protein"  
FT mutation replace(847, A)  
FT /\*tag= b  
XX EP1077260-A1.  
PN 21-FEB-2001.  
XX 13-AUG-1999; 99BP-00115161.  
XX 13-AUG-1999; 99BP-00115161.  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX Nueesch J, Rommelaere J;  
XX WPI; 2001-212717/22.  
DR P-PSDB; AAY72704.  
XX Novel parvovirus non-structure protein variant, useful for treating  
PT tumoral diseases, has a shifted equilibrium between DNA replication and  
PT transcription activities, and cytotoxic activity.  
XX Claim 7; Page 11-14; 41pp; English.  
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein  
CC (NS1) variant (S283A). The invention relates to the variants of the  
CC parvovirus non-structure protein (NS1) having a shifted equilibrium  
CC between the DNA replication and transcription activities, and the  
CC cytotoxicity activity. These variants are useful as toxins for treating  
CC tumoural diseases. The variant DNAs are useful as vectors for gene  
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)  
XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;  
SQ  
Query Match 99.8%; Score 2015.8; DB 5; Length 2019;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGGCTGGAATGCTTACTCTGATGAACTTTGGGAGCAACCACTGTTAAAGGAAAA 60  
Db 1 ATGGCTGGAATGCTTACTCTGATGAACTTTGGGAGCAACCACTGTTAAAGGAAAA 60  
Qy 61 AGTAACGAGCAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120  
Db 61 AGTAACGAGCAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120  
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGACGAGCTGAAATCTTTACAA 180  
Db 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGACGAGCTGAAATCTTTACAA 180  
Qy 181 CGAGGAGCGGAACACTACTTTGGGACCAAGGACATGGAATGGGAACCAACAGTGGAT 240  
Db 181 CGAGGAGCGGAACACTACTTTGGGACCAAGGACATGGAATGGGAACCAACAGTGGAT 240  
Qy 241 GAAATGACCAAAAAAGCAAGTATTTCATTTTGGTTTAAATAATGTTTATTGAA 300  
Db 241 GAAATGACCAAAAAAGCAAGTATTTCATTTTGGTTTAAATAATGTTTATTGAA 300

Qy 301 GTGCTTAACACAAAGAAATATATTTCTCGTGATGTTAAATGGTTTGTGCAACATGAATGG 360  
Db 301 GTGCTTAACACAAAGAAATATATTTCTCGTGATGTTAAATGGTTTGTGCAACATGAATGG 360  
Qy 361 GGAAGAAGACCAAGGCTGGCACTGCGCATCTGTAATAATGGAGGAAAGGACTTTAGTCAAGCT 420  
Db 361 GGAAGAAGACCAAGGCTGGCACTGCGCATCTGTAATAATGGAGGAAAGGACTTTAGTCAAGCT 420  
Qy 421 CAAGGGAATATGGTGAGAGGCAACTAAATGTTTACTGAGCAGATGGTTGTTAAACAGCC 480  
Db 421 CAAGGGAATATGGTGAGAGGCAACTAAATGTTTACTGAGCAGATGGTTGTTAAACAGCC 480  
Qy 481 TGTATGTGCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
Db 481 TGTATGTGCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540  
Qy 541 GAGTGGGTACTCTTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT 600  
Db 541 GAGTGGGTACTCTTACTTACTTATAAGCATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT 600  
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTTTAACTAAAGAAATAAGCACTAGT 660  
Db 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTTTAACTAAAGAAATAAGCACTAGT 660  
Qy 661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTCACTCTGCTGGAAACTTAACCTTTTAA 720  
Db 661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTCACTCTGCTGGAAACTTAACCTTTTAA 720  
Qy 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
Db 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780  
Qy 781 GTTGAACACCAAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTTAAAAAA 840  
Db 781 GTTGAACACCAAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTTAAAAAA 840  
Qy 841 GAAGTTTCTATTAATAACTACACTTAAAGAGCTGTGCTAATAAGAGTAACCTACCAAGAG 900  
Db 841 GAAGTTGCTATTAATAACTACACTTAAAGAGCTGTGCTAATAAGAGTAACCTACCAAGAG 900  
Qy 901 GACTGGATGATGTCAGCAGCAGAGTGTACATTGAAATGATGGCTCAACAGGTGGAGAA 960  
Db 901 GACTGGATGATGTCAGCAGCAGAGTGTACATTGAAATGATGGCTCAACAGGTGGAGAA 960  
Qy 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
Db 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020  
Qy 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCNAACCTAACCACTTTTCACTGCCTGAC 1080  
Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCNAACCTAACCACTTTTCACTGCCTGAC 1080  
Qy 1081 ACAAGAACCTGCAAGATTTTGTCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT 1140  
Db 1081 ACAAGAACCTGCAAGATTTTGTCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT 1140  
Qy 1141 ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCATGACCA 1200  
Db 1141 ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCATGACCA 1200  
Qy 1201 GCACAGCAGGCAATCTATTTATGCAAGCCATAGCAAGCAGTGGCAATGTTGGT 1260  
Db 1201 GCACAGCAGGCAATCTATTTATGCAAGCCATAGCAAGCAGTGGCAATGTTGGT 1260  
Qy 1261 TGCTATAATCAGCAATGTAACCTTTTCCATTTAAATGACGTACCAACAGAACTTGATT 1320  
Db 1261 TGCTATAATCAGCAATGTAACCTTTTCCATTTAAATGACGTACCAACAGAACTTGATT 1320  
Qy 1321 TGGGTAGAAGAGCTGGTAACCTTTGACAGCAAGTAACCAAGTTTAAAGCCATTGCTCT 1380  
Db 1321 TGGGTAGAAGAGCTGGTAACCTTTGACAGCAAGTAACCAAGTTTAAAGCCATTGCTCT 1380  
Qy 1381 GGTCAAGCTATTGCAATTGATCAAAAAAGGAAAGGCGCAGCAAAACAGATTGNAACCAACCA 1440

Db 1381 GGTCAAAACTATTGCAATTGATCAAAAAAGGAAAGGCGAGCAAAACAGATTGAAACCAACCA 1440  
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGTGTCAGAATAGGCTGCGAAGAAAGACCA 1500  
Db 1441 GTCATCATGACCAAAATGAGAACATTACAGTGTGTCAGAATAGGCTGCGAAGAAAGACCA 1500  
Qy 1501 GAACACACTCAACCAATCAGAGACAGATGTTTAACTTAACTTAACTTAACTTAACTTAACT 1560  
Db 1501 GAACACACTCAACCAATCAGAGACAGATGTTTAACTTAACTTAACTTAACTTAACTTAACT 1560  
Qy 1561 GGTGACTTTGTTGGTTGACAAAAATGAATGGCCCATGATTTGTTGCTTGGTTGTTAAAG 1620  
Db 1561 GGTGACTTTGTTGGTTGACAAAAATGAATGGCCCATGATTTGTTGCTTGGTTGTTAAAG 1620  
Qy 1621 AATGGTTTACCAATCTACCTAGGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTTCTGTATTGG 1680  
Db 1621 AATGGTTTACCAATCTACCTAGGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTTCTGTATTGG 1680  
Qy 1681 TCAGAAAACTGGGGGAGGCAAAAGGTGCGCAACTTCCTATAAATTTACTAGGTTGCGCAGCG 1740  
Db 1681 TCAGAAAACTGGGGGAGGCAAAAGGTGCGCAACTTCCTATAAATTTACTAGGTTGCGCAGCG 1740  
Qy 1741 TCACCAATTCACGACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Db 1741 TCACCAATTCACGACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800  
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGGACACACCAATATCTCTGTATTGG 1860  
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGGACACACCAATATCTCTGTATTGG 1860  
Qy 1861 GGCACTGCAAGAAACCCAGAACTGCGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Db 1861 GGCACTGCAAGAAACCCAGAACTGCGGGAAGCTGGTTTCCAAAGCCTGCCAAGATGGTCAA 1920  
Qy 1921 CTGAGCCCACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCCGTTGCGGAACCG 1980  
Db 1921 CTGAGCCCACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCCGTTGCGGAACCG 1980  
Qy 1981 TTGAAGAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019  
Db 1981 TTGAAGAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

## RESULT 7

AAT15311

ID AAT15311 standard; DNA; 5049 BP.

AC AAT15311;

XX

DT 14-OCT-1996 (first entry)

XX

Non-attenuated canine parvovirus CPV-39 passage 5 DNA.

DE

XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;

KW

XX ss.

OS

XX Canine parvovirus.

XX

PN WO9614088-A1.

XX

PD 17-MAY-1996.

XX

PF 02-NOV-1995; 95WO-US014207.

XX

PR 08-NOV-1994; 94US-00336345.

XX

(CORR ) CORNELL RES FOUND INC.

PA

XX Parrish CR, Gruenberg A, Carmichael LE;

PI

XX WPI; 1996-251556/25.

DR

XX







Qy	1915	GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTGAGAGCGTCTTCGGTGCG	1974
Db	2181	GTGCAAGCGAGTCCGACGTGTCGGAATAGAGGAGCAGACCTGAGAGCCATCTTTACTTCT	
Qy	1975	GAACCGTTGAAGAAAGACTTCAGCGA	2000
Db	2241	GAACAATTGGAAGAGATTTTCGAGA	2266
RESULT 8			
AAT15312			
ID	AAT15312 standard; DNA; 5049 BP.		
XX	AAT15312;		
XX	14-OCT-1996 (first entry)		
XX	Attenuated canine parvovirus CPV-39 passage 60 DNA.		
DE	Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;		
XX	ss.		
KW	Canine parvovirus.		
KW	XX		
OS	XX		
XX	XX		
XX	XX		
PH	Key Location/Qualifiers		
FT	misc_feature	59	
FT		/*tag= a	
FT		/note= "A, C or T"	
FT	misc_feature	97	
FT		/*tag= b	
FT		/note= "A, G or T"	
FT	misc_feature	4745	
FT		/*tag= c	
FT		/note= "A G or C"	
FT	misc_feature	4881	
FT		/*tag= d	
FT		/note= "A G or T"	
XX	XX		
PN	WO9614088-A1.		
XX	XX		
PD	17-MAY-1996.		
XX	XX		
XX	02-NOV-1995; 95WO-US014207.		
XX	XX		
PR	08-NOV-1994; 94US-00336345.		
XX	XX		
PA	(CORR ) CORNELL RES FOUND INC.		
XX	XX		
PI	Parrish CR, Gruenberg A, Carmichael LE;		
XX	XX		
DR	WPI; 1996-251556/25.		
XX	XX		
PT	Attenuated CPV strains contg. up to 4 mutation (s) relative to control		
PT	virus - useful as a veterinary vaccine against CPV disease in animals,		
PT	such as wild or domestic dogs.		
XX	XX		
PS	Claim 2; Page 24-27; 42pp; English.		
XX	XX		
CC	This viral DNA is isolated from an attenuated CPV. The DNA is preferably		
CC	derived from vB1440. The DNA is cloned into a vector which is used to		
CC	transfect a host cell. The vector used is preferably pGEM3Z or pGEM5Z.		
CC	The host cells to be transfected are selected from Norden Laboratory		
CC	feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or		
CC	canine A72 cells		
XX	XX		
SQ	Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;		
Query Match 51.5%; Score 1040.4; DB 2; Length 5049;			
Best Local Similarity 71.4%; Pred. No. 8.9e-270; Indels 18; Gaps 4;			
Matches 143; Conservative 0; Mismatches 556;			
Qy	1 ATGCGTGAAGTCTACTCTGATGAAGTTTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60		

Db 1350 GCAATTTCTAGACATGTCAAATTTTATAGAAATGCACGGATGGAAATTTGGAATTAAGTTTGT 1409  
QY 1135 CATGCTATTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATACTGTTTATTTTCAT 1194  
Db 1410 CAGCTATTAGCATGTGTTTAAATAGACAGAGTGGTAAAGAAATACAGTCTCTTTTCAT 1469  
QY 1195 GGACAGCCAGCAGACGCAAACTATTATTGCAAGCCATAGCAACAAGCAGTTGGCAAT 1254  
Db 1470 GGACAGCAAGTACAGGAAATCTATCTGCTCAAGCCATAGCAACAAGCTGTGGGTAAT 1529  
QY 1255 GTTGGTGTCTATTAATGCGCAATGAACTTTCCATTTTAATGACGTGTACCAACAGAAC 1314  
Db 1530 GTTGGTGTCTATTAATGCGCAATGAAATTTTCCATTTTAATGACGTGTACCAATAAAAT 1589  
QY 1315 TTGATTTGGGTAGAGAGCTGTGTAATTTGGGACAGCAAGTAAACAGTTTAAAGCCATT 1374  
Db 1590 TTAATTTGGATTGAGAGCTGTGTAATTTGGTCAACAGTTAATCAATTTAAAGCAATC 1649  
QY 1375 TGCTCTGGTCAAGCTATTGCGATTGATCAAAAGGAAAGGCGAGCAACAGATTTGAACCA 1434  
Db 1650 TGTCTGGACAAACAATTAGATTGATCAAAAGGTAAGGAAGTAAGCAAAATTTGAACCA 1709  
QY 1435 ACACCAAGTATCATGACCAACAAATGAGACATTTACAGTGTGAGATAGCGTCCGAGAA 1494  
Db 1710 ACTCCAGTAATTATGACAACTAATGAAATATTAACAATTTGAGAAATTTGGATGTGAAGAA 1769  
QY 1495 AGACCAAGACACACTCAACCAATCAGACAGCAAGATGCTTAACATTTCACTAACACATACC 1554  
Db 1770 AGACCTGAACATACACACCAATTAAGAGACAGAAATGTTGACATTTAGTATGTAG 1829  
QY 1555 TTGCTCTGGTCACTTTGTTGGTTGATGCAAAATGAAATGAGCCCAATTTGTTGCTGGTTG 1614  
Db 1830 CTTCCAGGAGACTTTGTTGGTTGATGAAAGAAATGGCTTTAATATGTGCATGGTTA 1889  
QY 1615 GTAAGATGTTTACCAATCTACATGCGCAAGTACTGTCTAAATGGGCAAGTTTCT 1674  
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTTAACATACATCATTTGGGGAAGTACCA 1949  
QY 1675 GATTTGGTTCAGAAACTGGGGGAGCCAAAGGTGCAACTCCTATAAATTTACTAGGTTGCG 1734  
Db 1950 GAATGGATGAAACTGGGGGAGCCTTAATATACAGAGGTTAATTTCAACAGGTTGC 2009  
QY 1735 GCAGCTCACCATTCAGACACCGAAAGTACGCTCTCAGCCAGAACATGACACTAAT 1794  
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCCTCAGAGTCAAGACCAAGTTCTAAT 2066  
QY 1795 CCACTTGCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCT 1854  
Db 2067 CCTCTGACTCCGGACGTAAGTGGACCTTGCACCTGGAACCGTGGAGTACTCCAGATACGCT 2126  
QY 1855 GTTGGCGGCACTGCAGAAACCCAGAACACATGCGGGAAGCTGGTTCCAAAGCCTGCCAAGAT 1914  
Db 2127 ATTGCAGAACTGCNAATCAACATCAACCAACTTGGCGTTACTCA-----CAAAGAC 2180  
QY 1915 GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTCGGTGCG 1974  
Db 2181 GTGCAAGCGAGTCCGACGTGGTCCGAAATAGAGAGCAGCCTGAGAGCCATCTTTACTTCT 2240  
QY 1975 GAACCGTTGAAGAAAGCTTCAGGA 2000  
Db 2241 GAACAATTTGAAGAAGATTTTCGAGA 2266

RESULT 9  
AAT88324  
ID AAT88324 standard; DNA; 5049 BP.

AC AAT88324;

XX AC

DT 17-OCT-2003 (revised)

DT 21-MAY-1998 (first entry)

XX Attenuated canine parvovirus (vBI440) genomic DNA.

XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;  
KW feline panleukopenia virus; mink enteritis virus; infection; ds.  
XX  
OS Canine parvovirus; vBI440 (ATCC VR 2489) .  
XX

Key Location/Qualifiers  
FH mutation 59  
FT /tag= C  
FT /note= "base 59 is G in CPV-39 (passage 5)"  
FT mutation 97  
FT /tag= d  
FT /note= "base 97 is C or T in CPV-39 (passage 5)"  
FT CDS 273..2279  
FT /tag= a  
FT /note= "NS1/NS2 coding region"  
FT CDS 2286..4541  
FT /tag= b  
FT /note= "VP1/VP2 coding region"  
FT mutation 4745  
FT /tag= e  
FT /note= "base 4745 is T in CPV-39 (passage 5)"  
FT mutation 4881  
FT /tag= f  
FT /note= "base 4881 is C in CPV-39 (passage 5)"

XX W09742972-A1.

PN 20-NOV-1997.

XX

XX

XX

PF 06-MAY-1997; 97WO-US007584.

XX

XX

PR 15-MAY-1996; 96US-00647655.

XX

XX (CORR ) CORNELL RES FOUND INC.

PA Parrish CR, Carmichael LE, Gruenberg A;

XX

PI WPI; 1998-008583/01.

XX

XX

XX

PT Canine parvovirus DNA carrying specific attenuating mutation(s) - used as

PT vaccines for protection against parvovirus and feline pan-leukopenia

PT virus infections.

XX

PS Example 8; Page 34-37; 60pp; English.

XX

XX This DNA sequence comprises an attenuated virus genome derived by serial

CC passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate

CC 39 in NUPK feline kidney host cells. The attenuated virus is designated

CC vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence

CC (see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations

CC are within the hairpin formed by the 3' terminal palindromic: the mutation

CC at nucleotide 59 introduces an A into a G-C rich region within the tip of

CC the hairpin, disrupting the base pairing in one of the 2 small internal

CC palindromes within that sequence; the thymine at nucleotide 97 is

CC adjacent to the mismatched bubble (flip-flop) sequence within the

CC palindromic. The DNA from attenuated CPV strains (see also AAT88321) is

CC used for the production of infectious molecular DNA clones, which, in

CC turn, can be transfected into cells to generate master stocks of the

CC virus. The attenuated viruses can be used in dogs as a vaccine to protect

CC against CPV disease, or more generally in cats and minks to protect

CC against feline panleukopenia virus and mink enteritis virus. The vaccines

CC protect against the currently prevalent CPV-2b type (and all extant

CC strains of types 2 and 2a), providing a long term immune response.

CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

XX

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 8.9e-270;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGCGTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCAACTGTTAAAGGAAAAA 60



KW Canine parvovirus; CPV; attenuation; vaccine; dogs;  
XX feline panleukopenia virus; mink enteritis virus; infection; ds.  
XX Canine parvovirus.

PH Key Location/Qualifiers

FT CDS 273...2279

FT FT /\*tag= a

FT FT /note= "NS1/NS2 coding region"

FT FT 2286...4541

FT FT /\*tag= b

FT FT /note= "VP1/VP2 coding region"

FT FT 4307

FT FT /\*tag= c

FT FT /note= "base 4307 is A in virulent CPV-39 (G in passage

FT FT 65 attenuated virus)"

FT FT 4358

FT FT /\*tag= d

FT FT /note= "base 4358 is C in virulent CPV-39 (T in passage

FT FT 65 attenuated virus)"

FT FT 4409

FT FT /\*tag= e

FT FT /note= "base 4409 is C in virulent CPV-39 (A in passage

FT FT 65 attenuated virus)"

FT FT 4477

FT FT /\*tag= f

FT FT /note= "base 4477 is G in virulent CPV-39 (T in passage

FT FT 65 attenuated virus)"

FT FT 4889

FT FT /\*tag= g

FT FT /note= "base 4889 is C in virulent CPV-39 (T in passage

FT FT 65 attenuated virus)"

FT FT 4973

FT FT /\*tag= h

FT FT /note= "base 4973 is C in virulent CPV-39 (T in passage

FT FT 65 attenuated virus)"

FT FT W09742972-A1.

XX XX 20-NOV-1997.

XX PF 06-MAY-1997; 97WO-US007584.

XX XX 15-MAY-1996; 96US-00647655.

XX XX (CORR ) CORNELL RES FOUND INC.

XX PI Parrish CR, Carmichael LE, Gruenberg A;

XX XX WPI; 1998-008583/01.

XX PT Canine parvovirus DNA carrying specific attenuating mutation(s) - used as  
XX PT vaccines for protection against parvovirus and feline pan-leukopenia  
XX PT virus infections.

XX XX Claim 1; Page: 60pp; English.

XX CC This DNA molecule encodes an attenuated canine parvovirus (CPV) genome.  
XX CC Attenuated viruses are obtained by serial passage of the virulent CPV  
XX CC type 2b isolate 39 in MLFV feline kidney host cells. They have one or  
XX CC more of the sequence alterations indicated in the sequence relative to  
XX CC the sequence of a control (5th passage) wild-type CPV-2b (see AAT88320).  
XX CC A claimed virus from the 65th passage (deposited as ATCC VR 2528)  
XX CC contains all 6 mutations. The DNA from attenuated CPV strains (see also  
XX CC AAT88324) is used for the production of infectious molecular DNA clones,  
XX CC which, in turn, can be transfected into cells to generate master stocks  
XX CC of the virus. The attenuated viruses can be used in dogs as a vaccine to  
XX CC protect against CPV disease, or more generally in cats and minks to  
XX CC protect against feline panleukopenia virus and mink enteritis virus. The  
XX CC vaccines protect against the currently prevalent CPV-2b type (and all  
XX CC extant strains of types 2 and 2a), providing a long term immune response.  
XX CC (NB. this sequence was created by adaptation of the wild-type CPV-2b  
XX CC sequence given in AAT88320)

XX SQ Sequence 5049 BP; 1788 A; 809 C; 1029 G; 1417 T; 0 U; 6 Other;  
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;  
Best Local Similarity 71.3%; Pred. No. 2.4e-269;  
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60  
Db ATGCTGGCAACCAGTATATCTGAGGAAGTTATGGAGGAGTAAATGGTTTAAAGAACAT 332  
QY 61 AGTAACGAGGAAGTGTCTCAATTTTGTATTTTAAATGAAATGTTCAACTGAATGAAAA 120  
Db GCAGAAAAATGAAGCATTTTCGTTTGTATTTTAAATGTGACAACTCCCACTAAATGGAAG 392  
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCAGGAGGACGAGCTGGAATCTTTTACAA 180  
Db 393 GATGTTCCGTGGAAACAACTATACCAACCAATTCAAATGAAGAGCTAACATCTTTAATT 452  
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAAACACA 234  
Db 453 AGAGGAGCACAAACAGCAATGGATCAACCGAAGAAAGAAATGACTGGGAATCGGAA 512  
QY 235 GTGGATGAATGACCAAAAAAGCAAGTATTTCATTTTGTATTTCTTTGGTTAAAAATGTTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAACTTTTGTATGCAATTAATAAAAATGTCTT 572  
QY 295 TTTTGAAGTGTCTTAACACAAAGAAATATATTTTCTGTGTGATGTTAATTTGGTTGTGCAACAT 354  
Db 573 TTTTGAAGTCTTTGTTTCTAAAAATATAGAACCAATGAATGTGTGTTTGTATTTCAACAT 632  
QY 355 GAATGGGAAAAAGACCAAGGCTGGCACTGCATGTACTAATTTGGAGGAAAGACTTTTAGT 414  
Db 633 GAATGGGAAAAAGATCAAGCTGGCATTTGTCATGTTTACTTCATAGTAAGAACTATACAA 692  
QY 415 CAAGCTCAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGAGTGGTTGTA 474  
Db 693 CAAGCAACTGGTAAATGGCTACGCAGACAAATGAATATGTATTTGGAGTAGATGGTTGGTG 752  
QY 475 ACAGCCTGTAATGTGCAACTAACACCAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAA 534  
Db 753 ACTCTTTGTCGTAACCTTAACCACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812  
QY 535 GACAATGAGTGGGTTACTCTTACTTACTTATAAGCATAAGCAAAACCAAAAGAACTATACC 594  
Db 813 GATAGTGAATGGTGACTATATTAACATACAGACATAAGCAAAACCAAAAGAACTATGTT 872  
QY 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATTAAGC 654  
Db 873 AAAATGGTTTCAATTTTGGAAATATGATAGCATATTTACTTTTAAACAAAGAAAAAATGTC 932  
QY 655 ACTAGTCCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTCTGGCTGGAATACTAAC 714  
Db 933 CACAT--GACAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGGTTGGAAATTTAAC 989  
QY 715 TTTTAAAGAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAGCACACTTTTACACTGAACAAATGAAACCA 1049  
QY 775 GAAACGGTTGAAACCAAGTAACCACTGGCAGGAAACTTAAGCGCGCAGAAATTCAAACT 834  
Db 1050 GAAACGGTTGAAACCAAGTACAGCAGCAGCAAGCAAGCAAGCGCGGAGAAATTCAAACT 1109  
QY 835 AAAAAAGAGTTTCTTATTAACCTACACTTTAAAGAGCTGGTGATCAAAAAGAGTAACCTCA 894  
Db 1110 AAAAGAAAGTGTCAATCAAAATGATCTTTGCGGAGCTTTGGTTAGTAAAGAGTAACATCA 1169  
QY 895 CCAGAGGACTGGATGATGATGACGACGACAGTATACATTTGAAATGATGCTCAACAGGT 954  
Db 1170 CCTGAAGACTGGATGATGTTTACAAACAGATAGTTATATTGAAATGATGGCACAACAGGA 1229  
QY 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAACCAA 1014

Db 1230 GGTGAAATCTTTTAAAAATACACTTGAAATTTGTACTTTGACTTTAGCAAGACAAA 1289  
Qy 1015 ACAGATTGACTTAATTTTAAAAAGCTGAACACCACTAACCAACTTTTCACTG 1074  
Db 1290 ACAGCAATTGAATTAATCTTTGAAAAAGCAGATAATACTAACTAACTTTGATCTT 1349  
Qy 1075 CCTGACACAGAACTGCAGAAATTTTGTCTTTTTCATGGCTGGAATATGTTAAAGTTGC 1134  
Db 1350 GCAAATTCAGAAATGTCAGAAATTTTGAAGATGACGAGTGAATTTGAAATTTGTT 1409  
Qy 1135 CATCTATTTCTGCTGTTTAAACAGACAGAGGCAAGAAATGCTTTTAAATTTTCAAT 1194  
Db 1410 CACGCTATGCAATGTTTAAATAGACAGAGTGTAAAGAAATACAGTCTTTTTCAT 1469  
Qy 1195 GGACAGCCAGCAGGCAAACTTATTATTCACAAGCCATPAGCAGCAAGTTGGCAAT 1254  
Db 1470 GGACAGCAAGTACAGGAAATCTATCTGCTCAAGCCATAGCAGCTGTGGTAAAT 1529  
Qy 1255 GTTGGTTGCTATATGAGCAATGTAACCTTTTCATTTAATGACTGTACCAAGAAC 1314  
Db 1530 GTTGGTTGTTATATGAGCAATGTAACCTTTTCATTTAATGACTGTACCAATAAAT 1589  
Qy 1315 TTGATTTGGGTAGAGAGCTGTTAACTTTTGACAGCAAGTAAACCACTTTAAAGCCATT 1374  
Db 1590 TTAATTTGGATTGAAGAGCTGTTAACTTTTGCTCAAGATTAATCAATTTAAAGCAATC 1649  
Qy 1375 TGCTCTGCTCAAGCTATTGCAATTTGATCAAAAAGAAAGGCAAGCAAGATTTGAACCA 1434  
Db 1650 TGTTCTGGACAACTTGAATTTGATCAAAAAGGTAAGGAAGTAAAGCAATTTGAACCA 1709  
Qy 1435 ACACAGTATCATGACCAAAATGAGAACATTAACAGTGTGACAAATAGGCTGCGAAGAA 1494  
Db 1710 ACTCCAGTAAATATGACAACTTAATGAAATTTTAACTTTGTGAGAAATTTGATGGAAGAA 1769  
Qy 1495 AGACCAAGACACACTCAACCAATCAGAGACAGATGCTTAACATTTCACTAAACATACC 1554  
Db 1770 AGACCTGACATACACAACTTAAGAGACAGAAATTTGACATTAATGATGATGTAAG 1829  
Qy 1555 TTGCTGCTGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1614  
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1889  
Qy 1615 GTAAAGAAATGGTTACCAATCTACATGCGAGCTGCTGCTAAATGGGGCAAGTTTCCT 1674  
Db 1890 GTTAAACATGGTTTGAATCAACCATGCTAACTATACATCATTTGGGAAAAAGTACCA 1949  
Qy 1675 GATTGGTCAGAAACTGGGGAGCCAAAGTGCCCACTCTCTATAAATTTACTAGTTGCG 1734  
Db 1950 GAATGGATGAAAACTGGGGAGCCCTAAATACAGAGGTATAAATTTACCAGGTTGC 2009  
Qy 1735 GCAGCTCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACATACT 1794  
Db 2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066  
Qy 1795 CCACTTCGATCGGATCTGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACCTCT 1854  
Db 2067 CCTCTGACTCCGAGCGTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCT 2126  
Qy 1855 GTTGGCGGCACTCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAGAT 1914  
Db 2127 ATTGCAGAACTGCAATCAACATCAACATCAACAACTTGGCGTTACTCA-----CAAAGAC 2180  
Qy 1915 GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCGGTGCG 1974  
Db 2181 GTGCAAGCGAGTCCGAGCTGTTCCGAATAGAGGACACCTGAGAGCCATCTTACTTCT 2240  
Qy 1975 GAACCGTTGAAGAAAGACTTTACGGA 2000  
Db 2241 GAACAAATTGGAAGAAAGATTTTCGAGA 2266

RESULT 11  
AAT88320

ID XX AAT88320 standard; DNA; 5049 BP.  
XX AAT88320;  
XX AC  
XX AC  
DT 17-OCT-2003 (revised)  
DT 21-MAY-1998 (first entry)  
XX  
XX Canine parvovirus 39 passage #5 (wild-type).  
XX  
XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;  
XX feline panleukopenia virus; mink enteritis virus; infection; ds.  
XX  
XX Canine parvovirus; type 2b isolate 39.  
XX  
XX  
FH Key Location/Qualifiers  
CDS 273..2279  
FT /\*tag= a  
FT /note= "NS1/NS2 coding region"  
FT 2286..4541  
FT /\*tag= b  
FT /note= "VP1/VP2 coding region"  
XX  
XX WO9742972-A1.  
XX  
XX 20-NOV-1997.  
XX  
XX 06-MAY-1997; 97WO-US007584.  
XX  
XX 15-MAY-1996; 96US-00647655.  
XX  
XX (CORR ) CORNELL RES FOUND INC.  
XX  
XX Parrish CR, Carmichael LE, Gruenberg A;  
XX WPI; 1998-008583/01.  
XX  
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as  
XX vaccines for protection against parvovirus and feline pan-leukopenia  
XX virus infections.  
XX  
XX Example 8; Page 37-40; 60pp; English.  
XX  
XX This DNA sequence comprises the genome of virulent canine parvovirus type  
XX 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NLFK feline  
XX kidney host cells. Further passaging has yielded attenuated virus vBI440  
XX (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus  
XX (see AAT88321) derived from the 65th passage (ATCC 2528). These  
XX respectively contain 4 and 6 mutations in comparison to the virulent 5th  
XX passage virus. The DNA from attenuated CPV-2b strains is used for the  
XX production of infectious molecular DNA clones, which, in turn, can be  
XX transfect into cells to generate master stocks of the virus. The  
XX attenuated viruses can be used in dogs as a vaccine to protect against  
XX CPV disease, or more generally in cats and minks to protect against  
XX feline panleukopenia virus and mink enteritis virus. The vaccines protect  
XX against the currently prevalent CPV-2b type (and all extant strains of  
XX types 2 and 2a), and provide a long term immune response. (Updated on 17-  
XX OCT-2003 to standardise OS field)  
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;  
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;  
Best Local Similarity 71.3%; Pred. No. 2.4e-269;  
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;  
Qy 1 ATGCTCGGAATGCTTACTCTGATGAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA 60  
Db 273 ATGCTCGGCAACCACTATATCTAGGAAGTTATCGGAGGAGTAAATGGTTTAAAGAAACAT 332  
Qy 61 AGTAACCAAGGAAGTGTCTCATTTGTTTAAATAATGMAAATGTTCAACTGAATGGAATA 120  
Db 333 GCAGAAATGAACATTTTCGTTGTTTAAATGTTGATGCAACGTCCTCAACTTAATGGAAG 392  
Qy 121 GATATCGGATGGAATAGTTACAAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA 180

[illegible]

Db	1470	GGACCAAGATGACGAGAAATCTATCATTTGCTCAAGCCATAGCACAAAGCTGTGGGTAA	1520
Qy	1255	GTGTGGTTGCTATAAATGCAGCAATATGTAACCTTTCCATTTAAATGACTGTACCAACAAGAAC	1314
Db	1530	GTGTGGTTGTTATAATGCAGCAAAATGTAATTTTCCATTTAAATGACTGTACCAATAAAAAT	1589
Qy	1315	TTGATTTGGGTAGAGAAGCTGTGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTT	1374
Db	1590	TTAATTTGGATTGAAGAAGCTGTGTAACTTTTGGTCAACAAGTTTAACTTTAAAGCAATC	1649
Qy	1375	TGCTCTGGTCTCAAGTATTCGATTTGATTCAAAAGGAAAAGGACGAGCAACAGATTGTAACCA	1434
Db	1650	TGTTCTTGGACAAACAATTTGATTCAAAAGGTTAAAGGAAAGTAGCAAAATTTGAACCA	1709
Qy	1435	ACACCAAGTCAATCATGACCAACAATGAGAACATTTACAGTGGTTCAGAAATAGGCTCGCAAGAA	1494
Db	1710	ACTCCAGTAAATATGACAACTAATGAAATTTTAACTTTGTGAGAAATGGATGTGAAGAA	1769
Qy	1495	AGACCAAGACACACTCAACCAATTCAGAGACAGAAATGCTTTAACTTCAATCTAAACACATACC	1554
Db	1770	AGACCTGAACATACACAACCAATAAGAGACAGAAATGTTGAACATTAAGTTAGTATGTAAAG	1829
Qy	1555	TTGCCTGGTGCATCTTGGTTTGGTTGTACAAAATGAATGGCCCAATGTTCTGCTTGGTTG	1614
Db	1830	CTTCCAGGAGACCTTGGTTTGGTTGTATAAGAAAGAAATGGCCCTTTAATATGTCATGGTTA	1889
Qy	1615	GTAAGAATGGTTTACCATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAGATTCCT	1674
Db	1890	GTTTAAACATGGTTTGTGAAATCAACCATGGCTTAATATACACATCATTTGGGAAAAGTACCA	1949
Qy	1675	GATTTGTCAGAAAATCTGGGCGGAGCCAAAGGTGCCAACTCCTATATAATTTACTAGGTTGG	1734
Db	1950	GAATGGGATGAAAACCTGGGCGGAGCCCTAAATAACAAGAAAGGTATAAATTCACCAGGTTGC	2009
Qy	1735	GCAGCTCACCATTCAGACACCGAAAGGTACGCCTCTCAGCCAGAACTATGCACATACT	1794
Db	2010	AAAGACT--TAGAGACAACAAGCGGCAAGCAATCTCAGAGTCAGACCAAGTTCTTAACT	2066
Qy	1795	CCACTTGCAATCGGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAAAATACTCCT	1854
Db	2067	CCTCTGACTCGGACGTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCCT	2126
Qy	1855	GTTCGGGCACTCGAGAAAACCCAGACACATGCGGGGAAGCTGGTTCCAAAGCCTGCCAAGAT	1914
Db	2127	ATTGCGAAAACCTGCAAAATCAACAATCAAAACCACTTTGGCGTTACTCA-----CAAAGAC	2180
Qy	1915	GGTCMACTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTCTTCGGTGGC	1974
Db	2181	GTGCAAGCGAGTCCGACGTGGTCGGAATATAGAGCGAGACCTTGAGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTTGAAGAAAGACTTCAGCGA	2000
Db	2241	GAACAATTTGAAGAGATTTTCGAGA	2266
RESULT 12			
ID	AAN40252 standard; DNA; 3524 BP.		
XX	AAN40252;		
AC	AAN40252;		
XX			
DT	24-OCT-2003 (revised)		
DT	12-JAN-1992 (first entry)		
XX			
XX			
DE	Sequence from the double-stranded replicative form DNA of porcine		
DE	parvovirus.		
XX			
KW	Protein envelope; immunogen; vaccine; antigen; epitope; ds.		
XX			
OS	Porcine parvovirus; NADL-2 virulent strain.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..2073	

```

FT      /*tag= a
FT      /note= "see AAP40306"
FT      CDS      2107..3522
FT      /*tag= b
FT      /note= "see AAP40675"
XX
PN      WO8402847-A.
XX
PD      02-AUG-1984.
XX
XX      19-JAN-1984; 84WO-US000063.
XX
XX      19-JAN-1983; 83US-00459203.
PR      06-JAN-1984; 84US-00567968.
XX
XX      (AMGE-) AMGEN.
XX
XX      Fox GM;
XX
XX      WPI; 1984-201354/32.
DR      P-PSDB; AAP40306, AAP40675.
XX
XX      Polypeptide obtd. by recombinant DNA methods - for vaccination against
XX      parvovirus infections in man and animals.
XX
XX      Claim 10; Table II, Page 33-49; 80pp; English.
XX
XX      The inventors claim an immunologically active polypeptide for the
XX      development of vaccinal immunity against parvovirus infection. Also
XX      claimed are DNA sequences wholly or partly duplicative of defined
XX      sequences. The polypeptides are used in vaccines for conferring
XX      protection against parvovirus infections in man and animals. (Updated on
XX      24-OCT-2003 to standardise OS field)
XX
XX      Query Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;
XX
XX      Query Match      44.6%; Score 901; DB 1; Length 3524;
XX      Best Local Similarity 69.7%; Pred. No. 3.2e-232;
XX      Matches 1296; Conservative 0; Mismatches 545; Indels 18; Gaps 5;
XX
QY      4 GCTGGAAATCCTTACTCTGATGAATTTTGGGAGCACCACTGGTTAAAGGAAAAGT 63
DB      13 GCGGGAACACTTACTCGGAAGAGGTACTAAAGACTACCAACTGGCTTCAAGATAATGCT 72
QY      64 AACGAGGAAGTGTCTCATTTGTTTTTAAATAATGTTCAACTGAATGAAAAGAT 123
DB      73 CAAGAAGACATCTCTTAATGTTTAAACACAAAAGTCAATCTAAATGGAAGAA 132
QY      124 ATCGGATGGAATAGTTACAAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACACGA 183
DB      133 ATTGCTTGGAAATAACTACAAACAAGATACACAGATGCGGAATGATAAACCTTACAAAGA 192
QY      184 GGAGCGGAAACTTGGGACCAAGAGGAGACATGGAATGGGAACCAAGTGGATGAA 243
DB      193 GGAGCAGAAACATCATGGGACCGGCAACAGACATGGAATGGGAATCAGAAATCGACAGC 252
QY      244 ATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTAAATTTGAAGTG 303
DB      253 CTCACAAAACGGCAAGTACTGATTTTGTGCTCTCTGTTTAAATAATGCTCTTTGAAGGT 312
QY      304 CTTTAAACAAAGAATATATTTCTGCTGATGTTAAATTTGTTTGTCAACATGAATGGGA 363
DB      313 ATATTGCAAAAGAACCTTAAGTCCAAAGTACTGCTACTGTTCTTACAGCATGAACATGGT 372
QY      364 AAGACCAAGCGTGGCACTGCCATGTAATTAATTTGGAGAAAGCACTTTAGTCAAGCTCAA 423
DB      373 CAAGATACCTGGCTATCACTGCCATGTACTAGTGGTGGAAAGGCTTACAAACAAGCAATG 432
QY      424 GGGAAATGTTGGAGGCACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCTGT 483
DB      433 GGAATAATGGTTACGAAACCAATTAACAAATTTATGGAGTAGATGGTTGAATATCAATGC 492
QY      484 AATGTGCAACTAACACCAGCTGAAAGAATTTAAACTAAGAGAAATAGCAGAGAACAATGAG 543

```

---

```

493 AAAGTACCTCTAACACCAGTTGAAAGAAATAAAATTAAGGAAATTTAGCAGAGGATGGTGAG 552
544 TGGGTACTCTTACTTATTAAGCAATAGCAAAACCAAAAAAGACTATATACCAAGTGTGTT 603
553 TGGGTATCGCTACTTAACCTACACTACAAACAACCTAAAAAACAATATACAAAAAGACT 612
604 CTTTTTGGAAAAATGATGCTTACTATTTTAACTAAATAAGAAAAAATAGCACTAGTCCA 663
613 CATTTTGGAAAAATGATGCTTACTTCTCTAAATAAAAAAAGAAAGACAACCT----- 666
664 CCAAGAGACGGAGCTATTTCTTACAGTACTCTGGCTGGGAAACTTCTTTTAAAAA 723
667 GAAAGAGAGCATGGATATTATCTCAGCTCAGATTTCTGGCTTCATGACAAAATTTCTTAAAA 726
724 GAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGCGCCAGAAACGGTT 783
727 GAAGCGAGAGACACTTAGTCAGTCACCTATTTACTTGAAGCAATAAACCCTGAACTGTG 786
784 GAAACCAAGTAAACCACTGCGCAGGAAACTTAAGCGCGGAGAAATTAACACTTAAAAAGAA 843
787 GAAACAACGGTTACTACAGCTCAGGAAGTTCCCGAGGCGAGAAATACAAACAAAAAAGAA 846
844 GTTCTTATTAATACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAGAC 903
847 GTAAGCATAAATGCAATAAGAGACTTGGTTTAATAAAAAAGATGTACTAGCATAGAAGC 906
904 TGATCATGATGCGCAGCAGACAGTTTAAATGATGCTCAACCCAGGTGAGAGAAAC 963
907 TGGATGATGACAGATCCAGACAGTTATATAGAAATGATGCTCAACCCGAGGAGAGAAAT 966
964 CTGCTGAAAAATAACGCTAGAGATTTGTACATACTAAGCTAGCCAGAAACCAAAACAGCATTT 1023
967 TTAATCAAAAAATACACTAGAAATPAAACAATCTTACTCTAGCAAGAAACAAAAACAGCATAT 1026
1024 GACTTAATTTTAGAAAAAGCTGAAACCGAGCAAACTAACCAACTTTTCACTGCCTGACACA 1083
1027 GACTTAATTTTGAAGGCAAAACCAAGCATGCTACCAACATTTTAATATTTAGCAATACA 1086
1084 AGAACCTGCGAGAAATTTTGTCTTTCATGCTGGAATCATGTTTAAAGTTTCCCACTGCTATT 1143
1087 AGAACATGTAATAATTTTCAAGTGCACAATTTGGAATCATATTAAGTCTGCCATGCTATA 1146
1144 TGCTGTGT-TTTAAACAGACAAGGAGGCAAAAGAAATA-----CTGTTTTTATTTCAATGGA 1197
1147 ACTTGTGCTGAAACAGACAAGGAGGAGAAAGAAATACAAATTTCTATTCTATGTCATGGG 1206
1198 CAGCCAGCAGCAGGCAAAATCTATTATTGCAAAAGCATAGCAACAGCAGTTGGCAATGTT 1257
1207 CCAGCATCAACAGGAAAAAGTATAATTGCTCAACACATTTGCAAACTTAGTTGGTAAATGTT 1266
1258 GGTGCTATAATGCGCAATGTAAACTTTTCCATTTTAAAGTGTACCAACAGAACTTG 1317
1267 GGTGCTATAATGCGCAATGTAAACTTTTCCATTTTAAAGTGTACCAATTAATAAACTTA 1326
1318 ATTTGGGTAGAAAGCTGTTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTCG 1377
1327 ATATGATTTGAAGAAGCAGGAAACTTCTTAACCAAGTAAACCAATTTCAAAGCCATATGT 1386
1378 TCTGGTCAAGCTATTTCGCAATTTGATCAAAAAGGAAAAAGGAGCAAAACAGATTTGAACCAACA 1437
1387 TCAGGTCAAACTATTAGAAATGACCAAAAAGGTAAGGAAAGCAAAACAAATTTGAACCAACT 1446
1438 CAGTATCATGACCAAAATGAGAACATTTACAGTGTGTCAGATAGCTGCGGAGAAAGA 1497
1447 CTTGTAATTAATGACTACAAATGAAGACATAAAGTTAAAGTTAGAAATGAGATCGGAGGAAAGA 1506
1498 CCAGAACACACTCAACCAATTCAGAGACAGATGCTTAAACATTTCTAAACACATACCTTG 1557
1507 CCAGAACATACACCAATTAAGAGACAGATGTTTAAACATACACCTTACCAAGAAACTG 1566
1558 CTTGTTGACTTTGGTTGTTGTCACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTA 1617

```

Db 1567 CCAGGTGATTTGGACTTTTAGAAGAACTGAATGGCCACTAATATGTGCTGGTTGGTA 1626  
Qy 1618 AAGAATGTTTACCAATCTACCATGGCAAGTACTGTCTAAATGGGCAAGTTCTGTGAT 1677  
Db 1627 AAGAAAGTTTACCAAGCAAACTGGCTAGTATATGCATCATTTGGGAAATGTACCTGAT 1686  
Qy 1678 TGGTCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTCGGCA 1737  
Db 1687 TGGTCAGAAAATTTGGAGGAGCCAAAATGCATTTCCCAATAAATACACCA---ACAGAC 1743  
Qy 1738 CGCTCACCATTACGACACCGAAAGTACGCCCTCTAGCCAGAACATATGCACTAATCCCA 1797  
Db 1744 TCTCAGATTTCCACATCAGTGAATACTTCGCCAGCGGACATCACTACGACGCAACTCCA 1803  
Qy 1798 CTTGCATCGATCTCGA---GGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATATCTCC 1853  
Db 1804 ATACAGGAGGACCTGGATTTAGCTTTAGCCCTTGGAGCGGTGGAGCCGCAACACACC 1862

RESULT 13  
ABQ95626  
ID ABQ95626 standard; DNA; 374 BP.  
AC ABQ95626;  
XX  
DT 28-OCT-2002 (first entry)  
XX  
DE Tumour suppression-related oligonucleotide #1277.  
XX  
KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
KW viral infection; cell degeneration disease; neurodegeneration; ds;  
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.  
XX  
OS Homo sapiens.  
XX  
PN FR2819824-A1.  
XX  
PD 26-JUL-2002.  
XX  
PF 23-JAN-2001; 2001FR-00000899.  
XX  
PR 23-JAN-2001; 2001FR-00000899.  
XX  
PA (MOLE-) MOLECULAR ENGINES LAB SA.  
XX  
PI Telerman A, Anson R, Tuijnder M, Susini L;  
XX  
DR WPI; 2002-610803/66.  
XX  
PT New nucleic acid implicated e.g. in tumor suppression, useful for  
PT diagnosis of tumors, viral infection and cellular degeneration and for  
PT drug screening.  
XX  
PS Claim 1; Page 360; 623pp; French.  
XX  
CC The present invention relates to novel human nucleic acid sequences (I).  
CC The present sequence is one such nucleic acid sequence. Expression of (I)  
CC are implicated in tumour suppression or reversion and apoptosis and viral  
CC resistance. (I) are useful as probes or primers for detecting,  
CC identifying, measuring and/or amplifying nucleic acid sequences, as  
CC antisense reagents and for recombinant production of polypeptides. (I),  
CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
CC these vectors and antibodies (Ab) against (II) are all useful for  
CC treatment/prevention of viral, tumour and cell degeneration diseases  
CC (especially neurodegeneration, such as Alzheimer's disease and  
CC schizophrenia). Analysing the expression of (I) is also useful for  
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
CC (I) are used for studying the aetiology of these diseases (also immune  
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
CC in the specification  
XX

SQ Sequence 374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;  
Query Match 15.2%; Score 307.4; DB 6; Length 374;  
Best Local Similarity 88.2%; Pred. No. 2e-72;  
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
Qy 174 TTTACAACGAGGCGGAGCACTTCTGGACCAAGCGAGGACATGGAATGGGAAACCCAC 233  
Db 1 TTTACAACGAGGCGGAGCACTTCTGGACCAAGCGAGGACATGGAATGGGAGAGCGC 60  
Qy 234 AGTGGATGAATGACCAAAAAGCAAGTATCTTTTGGATCTCTTGGTTAAATAATGTTT 293  
Db 61 AGTGGATGATGACCAAAAAGCAAGTATCTTTTGGATCTCTTGGTTAAATAATGTTT 120  
Qy 294 ATTTGAAGTGTCTTAACACAAAGAAATATATTTCTCGTGTGATGTTAAATTTGTTGCAACA 353  
Db 121 GTTTGAAGTGTCTGACACAAAGACATAGCTCTCTAGTAATGTTTCTGTCGACGA 180  
Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTAG 413  
Db 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTCATGTCTGATTTGGAGGCAAGGACTTTAG 240  
Qy 414 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGTCGACGATGTTTGT 473  
Db 241 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGTCGATGTTTGTGNN 300  
Qy 474 AACAGCTGTAAATGTGCAACTAACACAGCTGAAGAATTAACCTAAGAGAAATAGCAGA 533  
Db 301 GACTGCTGNAATGTNCACTAACACAGCTGAAGAATTAACCTGNNAGAAATAGCAGA 360  
Qy 534 AGCAATAGTGG 546  
Db 361 GGACAGTGANNNG 373

RESULT 14  
ABQ94779  
ID ABQ94779 standard; DNA; 421 BP.  
XX  
AC ABQ94779;  
XX  
DT 28-OCT-2002 (first entry)  
XX  
DE Tumour suppression-related oligonucleotide #430.  
XX  
KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
KW viral infection; cell degeneration disease; neurodegeneration; ds;  
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.  
XX  
OS Homo sapiens.  
XX  
PN FR2819824-A1.  
XX  
PD 26-JUL-2002.  
XX  
PF 23-JAN-2001; 2001FR-00000899.  
XX  
PR 23-JAN-2001; 2001FR-00000899.  
XX  
PA (MOLE-) MOLECULAR ENGINES LAB SA.  
XX  
PI Telerman A, Anson R, Tuijnder M, Susini L;  
XX  
DR WPI; 2002-610803/66.  
XX  
PT New nucleic acid implicated e.g. in tumor suppression, useful for  
PT diagnosis of tumors, viral infection and cellular degeneration and for  
PT drug screening.  
XX  
PS Claim 1; Page 141; 623pp; French.  
XX  
CC The present invention relates to novel human nucleic acid sequences (I).



CC The present sequence is one such nucleic acid sequence. Expression of (I)  
CC are implicated in tumour suppression or reversion and apoptosis and viral  
CC resistance. (I) are useful as probes or primers for detecting,  
CC identifying, measuring and/or amplifying nucleic acid sequences, as  
CC antisense reagents and for recombinant production of polypeptides. (I),  
CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
CC these vectors and antibodies (Ab) against (II) are all useful for  
CC treatment/prevention of viral, tumour and cell degeneration diseases  
CC (especially neurodegeneration, such as Alzheimer's disease and  
CC schizophrenia). Analysing the expression of (I) is also useful for  
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
CC (I) are used for studying the aetiology of these diseases (also immune  
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
XX in the specification  
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.2%; Score 285.8; DB 6; Length 421;  
Best Local Similarity 87.6%; Pred. No. 1.4e-66;  
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGAAACCA 232  
DB 59 CTTTACACGAGGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGAGAGCG 118  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGTTTAAATAATGTT 292  
DB 119 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTTTAAAGAGTGT 178  
QY 293 TATTTGAAGTGCTTAACACAAAGATATATTTCTGTGTATGTTAATTTGTTGCAAC 352  
DB 179 TGTTTGAAGTGCTCAGCACAAAGCAATAGTCTCTAGTATGTTACTTGGTTCGTGAGC 238  
QY 353 ATGAATGGGAAAGCAAGGCTGGCAGTCTGCTATGTTTCTTGGTTTAAAGAGTGT 412  
DB 239 ATGAATGGGAAAGCAAGGCTGGCAGTCTGCTATGTTTCTTGGTTTAAAGAGTGT 298  
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTTAAATGTTTACTTGGAGCAGATGTTGG 472  
DB 299 GTCAACCTCAAGGAAATGGTGAGAGGCAACTTAAATGTTTACTTGGAGTATGTTGG 358  
QY 473 TAACAGCTCTAATGTGCAACTAACACGCTGAAGAAATTAACCTAAGAGAAAT 527  
DB 359 TGACTGCTCTAATGTTTCAACTAACACGCTGAAGAAATTAACCTAAGAGAAAT 413

## RESULT 15

AB094724  
ID ABQ94724 standard; DNA; 423 BP.

AC ABQ94724;

XX 28-OCT-2002 (first entry)

DE Tumour suppression-related oligonucleotide #375.

XX Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;  
KW Tumour suppression; tumour reversion; apoptosis; viral resistance; human;  
KW viral infection; cell degeneration disease; neurodegeneration; ds;  
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS Homo sapiens.

XX FR2819824-A1.

XX 26-JUL-2002.

XX 23-JAN-2001; 2001FR-00000899.

XX 23-JAN-2001; 2001FR-00000899.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX

PI Telerman A, Anson R, Tuijnder M, Susini L;

XX WPI; 2002-610803/66.

DR New nucleic acid implicated e.g. in tumor suppression, useful for  
PT diagnosis of tumors, viral infection and cellular degeneration and for  
PT drug screening.

XX Claim 1; Page 125-126; 623pp; French.

XX The present invention relates to novel human nucleic acid sequences (I).  
XX The present sequence is one such nucleic acid sequence. Expression of (I)  
CC are implicated in tumour suppression or reversion and apoptosis and viral  
CC resistance. (I) are useful as probes or primers for detecting,  
CC identifying, measuring and/or amplifying nucleic acid sequences, as  
CC antisense reagents and for recombinant production of polypeptides. (I),  
CC polypeptides (II) encoded by (I), vector containing (I), cells containing  
CC these vectors and antibodies (Ab) against (II) are all useful for  
CC treatment/prevention of viral, tumour and cell degeneration diseases  
CC (especially neurodegeneration, such as Alzheimer's disease and  
CC schizophrenia). Analysing the expression of (I) is also useful for  
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying  
CC (I) are used for studying the aetiology of these diseases (also immune  
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1  
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown  
XX in the specification  
SQ Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.1%; Score 284.8; DB 6; Length 423;  
Best Local Similarity 89.2%; Pred. No. 2.7e-66;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGAAACCA 232  
DB 71 CTTTACACGAGGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGAGAGCG 130  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGTTTAAATAATGTT 292  
DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTTTAAAGAGTGT 190  
QY 293 TATTTGAAGTGCTTAACACAAAGATATATTTCTGTGTATGTTAATTTGTTGCAAC 352  
DB 191 TGTTTGAAGTGCTCAGCACAAAGCAATAGTCTCTAGTATGTTACTTGGTTCGTGAGC 250  
QY 353 ATGAATGGGAAAGCAAGGCTGGCAGTCTGCTATGTTAATTTGGAGGAAAGCACTTTA 412  
DB 251 ATGAATGGGAAAGCAAGGCTGGCAGTCTGCTATGTTAATTTGGAGGCAAGCACTTTA 310  
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTTAAATGTTTACTTGGAGCAGATGTTGG 472  
DB 311 GTCAACCTCAAGGAAATGGTGAGAGGCAACTTAAATGTTTACTTGGAGTATGTTGG 370  
QY 473 TAACAGCTCTAATGTGCAACTAACACGCTGAAGAAATTAACCTAAGAGAAAT 516  
DB 371 TGACTGCTCTAATGTTTCAACTAACACGCTGAAGAAATTAACCTAAGAGAAAT 414

Search completed: January 22, 2005, 11:50:01

Job time : 957.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds  
(without alignments)  
8723.907 Million cell updates/sec

Title: US-10-069-056-16

Perfect score: 2019

Sequence: 1 atggctggaatgcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	107.8	5.3	4680	1	US-08-254-358-1
6	107.8	5.3	4680	1	US-08-475-391-1
7	107.8	5.3	4680	2	US-08-709-609-1
8	107.8	5.3	4680	5	PCT-US95-07178-1
9	107.8	5.3	4681	4	US-08-807-802A-18
10	107.8	5.3	4910	2	US-08-331-384-2
11	107.8	5.3	4910	2	US-08-836-087-2
12	107.8	5.3	4910	3	US-09-246-320-2
13	107.8	5.3	4910	3	US-09-546-738-2
14	107.8	5.3	7214	4	US-09-438-268-1
15	107.8	5.3	7557	4	US-09-770-315-3
16	107.8	5.3	8151	4	US-09-438-268-2
17	107.8	5.3	8179	4	US-09-438-268-5
18	107.8	5.3	8698	4	US-09-770-315-2
19	103	5.1	939	4	US-09-532-594B-12
20	103	5.1	1197	4	US-09-532-594B-13
21	103	5.1	1611	4	US-09-532-594B-14
22	103	5.1	1872	4	US-09-532-594B-3
23	103	5.1	1872	4	US-09-532-594B-15
24	103	5.1	4767	4	US-09-532-594B-1
25	98.2	4.9	969	4	US-09-807-802A-10
26	98.2	4.9	1200	4	US-09-807-802A-8
27	98.2	4.9	1641	4	US-09-807-802A-6

Sequence 4, Appli  
Sequence 19, Appli  
Sequence 1, Appli  
Sequence 14, Appli  
Sequence 22, Appli  
Sequence 22, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 2813, Ap  
Sequence 14265, A  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 325, App  
Sequence 8976, Ap

#### ALIGNMENTS

RESULT 1  
US-08-336-345-1  
; Sequence 1, Application US/08336345  
; Patent No. 5814510  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,345  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30753  
; REFERENCE/DOCKET NUMBER: 7937-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Parvovirus  
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 1.2e-292;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGACCAACCACTGTTAAGGAAAAA 60

Db 273 ATCTCTGGCAACCCAGTATATCTGAGGAAGTTATGGAGGAGTAAATGGTTAAAGAAACAT 332  
Qy 61 AGTAAACGAGGAAGTGTCTCATTTGTTTAAATAAGAAATGTTCAACTGATGGAAGA 120  
Db 333 GCAGAAAATGAAGCAATTTTGGTTGTTTAAATGTGACAAAGTCCAACTAAATGGAAG 392  
Qy 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCCAGGAGGACGAGCTGAAATCTTTTACAA 180  
Db 393 GATGTTGCTGGAAACAACTATACAAACCAATTCAAAATGAAGAGCTAACATCTTTAAAT 452  
Qy 181 CGAGGAGCGGAACCTACTTTGGGACCAAGC-----GAGGACATGAAATGGGAAACCA 234  
Db 453 AGAGGAGCACAACAGCAATGGATCAACCGAAGAGAAATGACTGGGAATCGGA 512  
Qy 235 GTGATGAATGACCAAAAAGCAAGTATTCATTTTGTGTTGTTGTTTAAATAATGTTTA 294  
Db 513 GTTGATAGTCTCGCCAAAAGCAAGTACAAACTTTTGTGATTAATTAATAAATGTCTT 572  
Qy 295 TTTGAAGTGTCTTAACACAAAGAAATATATTTCTGTGTGATGTTAAATGGTTTGTGCAACAT 354  
Db 573 TTTGAAGTCTTTGTTCTTAAATATAGAACCAATGAATGTGTTGTTTATTCAACAT 632  
Qy 355 GAATGGGAAAGACCAAGCTGGCACTGCCATGTACTTAATTCGAGGAAGGACTTTAGT 414  
Db 633 GAATGGGAAAGATCAAGCTGGCAATGTCTATGTTTACTTCTATAGTAAAGACTTACAA 692  
Qy 415 CAAGCTCAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTA 474  
Db 693 CAAGCAACTGGTAAATGGCTACCGAGACAAATGAATGATTTGGAGTAGATGGTGGTG 752  
Qy 475 ACAGCTGTAAATGTGCAACTAAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534  
Db 753 ACTCTTTGTTCCGTAAACTTTAACACCAACTGAAAGATTAAGCTCAGAGAAATTCGAGAA 812  
Qy 535 GACAAATGAGTGGGTACTCTACTTACTTATAGCAATAGCAACCAACCAAGAAAGTATACC 594  
Db 813 GATAGTCAATGGGTGACTATATTAACATACAGACATAAGCAAAACAAAGAAAGAACTATGTT 872  
Qy 595 AAGTGTCTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAAGAAATTAAGC 654  
Db 873 AAAATGGTTCATTTTGGAAATATGATAGCAATATCTTTTAAAGAAAGAAATTTGTC 932  
Qy 655 ACTAGTCCACCAAGAGACGGAGCTATTTTCTTAGCAGTGAATCTGTGCTGGAAACCTAAC 714  
Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGGTTGGAATTTAAC 989  
Qy 715 TTTTAAAGAGCGGAGCGCCATCTAGTGAGCAACTATACACTGATGATGATGCGGCCA 774  
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGAACCA 1049  
Qy 775 GAAACGGTTGAAACCAACAGTAACCACTCGCGAGAAACTAAGCGCGGAGAAATTCAAACT 834  
Db 1050 GAAACCGTTGAAACCAACAGTACAGACAGACAGAAACAAAGCGCGGAGAAATTCAAACT 1109  
Qy 835 AAAAAGAGAGTTCTTATTAATAACTTACACTTAAAGAGTGGTGATATAAAGAGTAACCTCA 894  
Db 1110 AAAAAGGAAGTGTCAATCAATGTACTTTGCGGACTTGGTTAGTAAAGAGTAACATCA 1169  
Qy 895 CCAGAGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
Db 1170 CTTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229  
Qy 955 GGAGAAACCTGCTGAAATAACGCTAGAGATTTGTACACTAATCTTAGCCAGAACCAAA 1014  
Db 1230 GGTGAAATCTTTTAAATAATACACTTGAAATTTGTTGACTTTTGTAGCAAGAAACAAA 1289  
Qy 1015 ACAGATTTGACTTAAATTTTAAAGAAAGCTGAAACCGACCAACTAACCACTTTTCACTG 1074  
Db 1290 ACAGCATTTGAATTAATACTTTGAAAGAGCAGATATTAATACTAACTAACTTTTGTGCTT 1349  
Qy 1075 CTTGACAAAGAACTGAGAAATTTTCTTTTCTGCTGCAACTATGTTTAAAGTTTGC 1134

## RESULT 2

US-08-336-345-2  
; Sequence 2, Application US/08336345  
; Patent No. 5814510  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland B.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2

Db 1350 GCAAAATCTAGAAATGTCATAATTTTTAGAAATGCACGATGGAATTTGGATTAAAGTTGT 1409  
Qy 1135 CATGCTATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAAGAAATACCTGTTTTATTTCAT 1194  
Db 1410 CACGCTATAGCATGTGTTTTTAAATAGACAAGGTGTAAGAAATACAGTTCTTTTTCAT 1469  
Qy 1195 GGACGAGCAGACAGGCAAAATCTATTATTCGCAAAAGCCATAGCACAGAGCTTGGCAAT 1254  
Db 1470 GGACGAGCAAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCACAGCTGTGGGTAA 1529  
Qy 1255 GTTGTGTGCTAATAATGCAGCCAAATGTAACCTTTTCATTTTAATGACTGTACCAACAGAAC 1314  
Db 1530 GTTGTGTGTTAATAATGCAGCAATGTAATTTTTCATTTAATGACTGTACCAATAAATA 1589  
Qy 1315 TTGATTTGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACAGCTTTTAAAGCCATT 1374  
Db 1590 TTAATTTGGATTGAAGAGCTGGTAACTTTGGTCAACAAAGTTAACTCAATTTTAAAGCAATC 1649  
Qy 1375 TGCTCTGTCAAGCTATTTCGCAATTTGATCAAAAGGAAAGGAGGAGGAGGAGGAGGAGGAGG 1434  
Db 1650 TGTTCTGGACAAACAAATTAGAATTTGATCAAAAGGTAAGGAAAGTAAGCAAAATTTGAACCA 1709  
Qy 1435 ACACGAGTCACTCATGACCAACAAATGAGAACATTAACAGTGGTCAGAAATAGGCTGCGAAGAA 1494  
Db 1710 ACTCCAGTAAATTTGACAACTTAATGAAATATAACAATTTGTGAGAAATTTGGATGTGAAGAA 1769  
Qy 1495 AGACGAGAAACACACTCAACCAATCAGACAGACAGAAATGCTTTAACTTCACTTAACATACC 1554  
Db 1770 AGACCTGAACTACACCAACCAATAAGAGACAGAAATGTTGAACATTTAAGTTAGTATGTAG 1829  
Qy 1555 TTGCTCTGTGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1614  
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1889  
Qy 1615 GTAAAGAAATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTPAAATGGGGCAAGTTCTCT 1674  
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACACATCAATTTGGGGAAGATACCA 1949  
Qy 1675 GATTGTGCAAGAACTGGGCGGAGCAAGGTGCCAACTCTCTATAAATTTACTAGTTCTG 1734  
Db 1950 GAATGGATGAAACTGGGCGGAGCTTAAATACAAAGAGGTATAAATTCACAGGTTGC 2009  
Qy 1735 GCAGCTCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794  
Db 2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGCCAAAGTTCTAACT 2066  
Qy 1795 CCATTTGCTATCGGATCTCGAGGACCTGCTTTAGAGCTTTGAGCAGCACCAAAATCTCCT 1854  
Db 2067 CCTCTGACTCCGACCGTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCT 2126  
Qy 1855 GTTGGGGCACTGCAGAAACCCAGAACACTGCGGAGAGCTGTTCCAAAGCTGCCAAGAT 1914  
Db 2127 ATTGCAGAACTGCAAAATCAACAACTCAAACTTTGGCGTTACTCA-----CAAAGAC 2180  
Qy 1915 GGTCAACTGAGCCCAACTTTGTTGAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGG 1974  
Db 2181 GTCGAAGCGAGTCCGAGCTGGTCCGAATAGAGGAGGAGCTGAGAGCCATCTTTACTTCT 2240  
Qy 1975 GAACCGTTGAAGAAAGACTTTCAAGCA 2000  
Db 2241 GAACAATTTGAAGAAAGATTTTCGAGA 2266



```
QY 1615 GTAAAGAAATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCGCAAGTTCTCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACACATCATTTGGGGAAAGATACCA 1949
QY 1675 GATTGCTGAGAAATCTGGGCGGAGCAAAAGGTGCCAACTCCCTATATAAATTTACTAGTTGCG 1734
Db 1950 GAATGGATGAAATCTGGGCGGAGCTAAATACAAAGAGGTATAAATTCACCGGTTGC 2009
QY 1735 GCAGGCTCACCATTACGACACCGAAAGTACGCTCTCAGCGAGAACTATGCTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCCTCAGAGTCAAGACCAAGTTCTTAAT 2066
QY 1795 CCATTTGCTATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATACTCTCT 1854
Db 2067 CCTCTGACTCCGAGCTAGTGGACCTTGCACTGNAACCGTGGAGTACTCCAGATACGCT 2126
QY 1855 GTTTCGGGCACTCGAGAAACCCAGAACACTGTGGGGAAGCTGTTCCAAAGCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTCAAAATCAACATCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180
QY 1915 GGTCAACTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGG 1974
Db 2181 GTCCNAGCGAGTCCGACGTGGTCCGAAATAGAGCGAGACTGAGAGCCATCTTTACTTCT 2240
QY 1975 GAACCGTTGAAGAAAGACTTCAGCGA 2000
Db 2241 GAACAAATTGAAGAAGATTTTCGAGA 2266
```

RESULT 3

```
US-08-647-655-1
; Sequence 1, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
```

US-08-647-655-1

```
Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.2e-292;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGAAAGAAA 60
Db 273 ATGTCTGGCAACCACTATATCTGAGGAAGTTATGGAGGAGTAAATTTGGTTAAAGAAACAT 332
QY 61 AGTAACCCAGGAAGTGTCTCTATTCTTTTAAAAATGAAATGTTTCAACTGAATGAAAA 120
Db 333 GCAGAAATGAAGCAATTTCTGTTTGTATTAATGTGCAACGTCCAACTAATGGAAG 392
QY 121 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGAGTGAATCTTTTACAA 180
Db 393 GATGTTGCGTGAACCAACTATACCAACCAATTTCAAAATGAAGAGCTAACATCTTTAAT 452
QY 181 CGAGGAGCGGAACCTACTTGGGACCAAGC-----GAGGACATGGAATGGGAACCA 234
Db 453 AGAGGAGCAACAACAGCAATGGATCAACCGGAAGAGAAATGGACTGGGAATCGGAA 512
QY 235 GTGGATGAATGACCAAAAAAGCAAGTATTTCATTTTGTGTTTCTTTGTTAAAAATGTTTA 294
Db 513 GTTGATGCTCGCCAAAAAGCAAGTACAACTTTTGTGCAATTAATTAATAAATGCTCT 572
QY 295 TTTGAAGTCTTTAACACAAAGAAATATATTTCTGTTGATGTTAAATTTGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTTCTAAAAATATAGAAACCAAAATGAATGTGTGTTTGTGTTTCAACAT 632
QY 355 GAATGGGGAAGAACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGAGCTTTTACT 414
Db 633 GAATGGGGAAGATCAAGGCTGGCAATTTGTCAATTTTCTTCACTTCTCATAGTAAGAACCT 474
QY 415 CAAGCTCAAGGAAATGGTGGGAAGGCACTAAATGTTTCTTGGAGCAGATGGTTGGTA 474
Db 693 CAAGCACTGTAATGGCTACGACACAAATGAATATGTTTGGAGTAGATGGTTGGTG 752
QY 475 ACAGCCTGTAATGTGCAACTAACACAGCTGAAAGAAATTAATAAGTAAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAAACTTAAACCACTGAAAGAAATTAAGCTCAGAGAAATTTGAGAA 812
QY 535 GACAACTGAGTGGTTTACTCTCTACTTATTAAGCATAAGCAACCAACCAAGAAAGCTATACC 594
Db 813 GATAGTAATGGTGACTATATTAACATACAGACATAGCAACCAACCAAGAAAGCTATGTT 872
QY 595 AAGTGTGTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAAAGAAATTAAGC 654
Db 873 AAAATGTTCAITTTTGGAAATATGATGACATATTTACTTTTAAACAAAGAAAGAAATTTGTC 932
QY 655 ACTAGTCCCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGCTGGTGGAAAGCTTAAC 714
Db 933 CACAT---GACAAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGTTGGTAAATTTAAC 989
QY 715 TTTTAAAGAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAGACACTTTTACTGACAAATGAAACCA 1049
QY 775 GAAACGGTTGAAACCCACAGTAACCTGCGCAGGAAACTAAGCGCGCAGAAATTTCAAACT 834
Db 1050 GAAACCGTTGAAACCCACAGTGACGACAGCAAGCAACAAAGCGCGGAGAAATTTCAAACT 1109
QY 835 AAAAAAGAGTTTCTATTAATAACTACACTTAAAGAGCTGGTGATTAAGAGTAACCTCA 894
Db 1110 AAAAAAGAGTGTCAATCAAAATGTTTTCGCGGACTTTGTTAGTAAAGAGTAACATCA 1169
QY 895 CCAGAGGACTGGATGATGATGACGACAGCAGTATACATTTGAAATGATGGCTCAACAGGT 954
Db 1170 CTTGAAGACTGGATGATGTTTACCAACAGATAGTTTATTTGAATGATGGCACAACCGGA 1229
QY 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACTAACTAGCTAGCAGAACCAAA 1014
Db 1230 GGTGAAAACTCTTTTAAAAAATAACACTTGAAATTTGTACTTTGACTTTTAGCAAGAACAAA 1289
```







```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match 5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0

QY 1168 GGCAAAAGAAATACTGTTTATTTATTCATGACAGCAGCAGCAGGCAAAATCTATTATTCGA 1227
DB 1293 GGCAGAGGAACACCAATCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1352
QY 1228 CAAAGCCATAGCACAAGCAGTGGCAATGTTGCTGCTATTAATGACGCCAATGTAAACTTT 1287
DB 1353 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGGTAAACTGGACCAATGAGAACTTT 1412
QY 1288 CATTATTAATGACTGTACCAAGAACTTGTATTTGGGTGAGAAAGCTGGTAACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTGCACAGATGGTGTCTGTGGGAGGAGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGTGTCAGCTATTTCGCATTGATCAAAA 1407
DB 1473 AAGTCTGGAGTCGCCCAAGCCATTTCTCGAGGAAGCAAGTGGCGGTGGACCAAGAA 1532
QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACCAAGTCATGACCAACAAATGAGAACTT 1467
DB 1533 TGCAAGTCTCGGCCAGATAGACCGACTCCCGTGATCGTCACTCCACCAACATG 1592
QY 1468 ACAGTGGTCAGAAATAGGCTGGGAAGAAAGACAGAACACACTCAACCAATCAGAGACAGA 1527
DB 1593 TGCGCGGTGATTGACGGGAACCTCAAGACCTTCGAACACCAAGCAGCCGTTGCAAGACGG 1652
QY 1528 ATGCTTAAATCATCTACACATACCTTGCCTGCTGCTGCTGTTGTTGTTGCAAAAT 1587
DB 1653 ATGTTCAATTTGAATCACTCCCGCGCTGATGATGATGATGATGATGATGATGATGATG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 6
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match 5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0

QY 1168 GGCAAAAGAAATACTGTTTATTTATTCATGACAGCAGCAGCAGGCAAAATCTATTATTCGA 1227
DB 1293 GGCAGAGGAACACCAATCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1352
QY 1228 CAAAGCCATAGCACAAGCAGTGGCAATGTTGCTGCTATTAATGACGCCAATGTAAACTTT 1287
DB 1353 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGGTAAACTGGACCAATGAGAACTTT 1412
QY 1288 CATTATTAATGACTGTACCAAGAACTTGTATTTGGGTGAGAAAGCTGGTAACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTGCACAGATGGTGTCTGTGGGAGGAGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGTGTCAGCTATTTCGCATTGATCAAAA 1407
DB 1473 AAGTCTGGAGTCGCCCAAGCCATTTCTCGAGGAAGCAAGTGGCGGTGGACCAAGAA 1532
QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACCAAGTCATGACCAACAAATGAGAACTT 1467
DB 1533 TGCAAGTCTCGGCCAGATAGACCGACTCCCGTGATCGTCACTCCACCAACATG 1592
QY 1468 ACAGTGGTCAGAAATAGGCTGGGAAGAAAGACAGAACACACTCAACCAATCAGAGACAGA 1527
DB 1593 TGCGCGGTGATTGACGGGAACCTCAAGACCTTCGAACACCAAGCAGCCGTTGCAAGACGG 1652
QY 1528 ATGCTTAAATCATCTACACATACCTTGCCTGCTGCTGCTGTTGTTGTTGCAAAAT 1587
DB 1653 ATGTTCAATTTGAATCACTCCCGCGCTGATGATGATGATGATGATGATGATGATGATG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 7
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.

```

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/709,609
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 585877sand, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 31975
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4680 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-709-609-1

Query Match 5.3%; Score 107.8; DB 2; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGCAGAGGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1227
DB 1293 GGCAGAGGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1352
QY 1228 CAAGCCATAGCAGCAAGAGTTGGCAATGTTGGTGTCTATTAATCGACCAATGTAACCTTT 1287
DB 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGGGTAACTGGACCAATGGAACCTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAGCAATGTTGGTGTCTATTAATCGACCAATGTAACCTTT 1347
DB 1413 CCCTTCAACGACTGTGTCAGCAAGATGTTGTTGGTGGAGGAGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGATTTAAAGCAATTTGCTCTGCTCAAGCTATTTCGCAATTGATCAAAA 1407
DB 1533 TGCAAGTCTCGGCGGAGTAGACCCGACTCCGCGTATCGTCACTCCACACCAATG 1592
QY 1468 ACAGTGTGAGAAATAGCTCGAAGAAAGACACACTCAACCAATCAGACAGAGA 1527
DB 1593 TGGCGCGTATTGACGGGAATCAACGACCTTCGAAACACAGCAGCGGTTGCAAGACCG 1652
QY 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGAATTTGGTTTGTGCAAAAAT 1587
DB 1653 ATGTTCAAAATTTGAATCAACCCCGCTCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

```
RESULT 8
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
```

```
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/07178
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Noland, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 31975
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4680 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
PCT-US95-07178-1
```

```
Query Match 5.3%; Score 107.8; DB 5; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGCAGAGGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1227
DB 1293 GGCAGAGGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1352
QY 1228 CAAGCCATAGCAGCAAGAGTTGGCAATGTTGGTGTCTATTAATCGACCAATGTAACCTTT 1287
DB 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGGGTAACTGGACCAATGGAACCTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAGCAATGTTGGTGTCTATTAATCGACCAATGTAACCTTT 1347
DB 1413 CCCTTCAACGACTGTGTCAGCAAGATGTTGTTGGTGGAGGAGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGATTTAAAGCAATTTGCTCTGCTCAAGCTATTTCGCAATTGATCAAAA 1407
DB 1473 AAGTCTCGTGGAGTAGACCCGACTCCGCGTATCGTCACTCCACACCAATG 1532
QY 1408 GGAAGAGGAGCAAAACAGATTTGAACCAACACAGATTCATCATGACCAAAATGAGAACAT 1467
DB 1533 TGCAAGTCTCGGCGGAGTAGACCCGACTCCGCGTATCGTCACTCCACACCAATG 1592
QY 1468 ACAGTGTGAGAAATAGCTCGAAGAAAGACACACTCAACCAATCAGACAGAGA 1527
DB 1593 TGGCGCGTATTGACGGGAATCAACGACCTTCGAAACACAGCAGCGGTTGCAAGACCG 1652
QY 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGAATTTGGTTTGTGCAAAAAT 1587
DB 1653 ATGTTCAAAATTTGAATCAACCCCGCTCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

```
RESULT 9
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
```

; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 18  
; LENGTH: 4681  
; TYPE: DNA  
; ORGANISM: AAV-2  
; US-09-807-802A-18

Query Match 5.3%; Score 107.8; DB 4; Length 4681;  
Best Local Similarity 53.4%; Pred. No. 8.6e-21;  
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 1168 GGCAGAGGAACTCTGTTTATTCATGGACCGCCAGCAGCAGGCAAAATCTATTATGCA 1227  
DB 1293 GGCAGAGGAACTCTGTTTATTCATGGACCGCCAGCAGCAGGCAAAATCTATTATGCA 1352  
QY 1228 CAAAGCCATAGCACAAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCAAATGTAACATTT 1287  
DB 1353 GAGGCCATAGCCCACTGTGCGCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1412  
QY 1288 CCATTAAATGACTGTACCAAGAACTTTGATTGGGTAGAAAGCTGTTAACTTTGGA 1347  
DB 1413 CCCTTCAACGACTGTGTCACAGATGTTGATCTGTTGGGAGGAGGAGATGACCGCC 1472  
QY 1348 CAGCAAGTAAACCAAGTTAAAGCCATTTGCTCTGCTCAAGCTATTTCGATTTGATCAAAA 1407  
DB 1473 AAGTCTGTGAGTGGCGCAAGCAATTTCTCGAGGAAGCAAGTGGCGGTGGACCAAGAA 1532  
QY 1408 GGAAGGAGCAGAAACAGATTGAACCAACACAGTCTATCATGACCAACAAATGAGACATT 1467  
DB 1533 TGCAAGTCTCGGCCAGATAGACCCGACTCCCGTGTATGTCACCTCCACCAACATG 1592  
QY 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527  
DB 1593 TGCGCGTGTATTCAGGGAACTCAACGACCTTCGAACACAGCAGCGTTGCAAGCCGG 1652  
QY 1528 ATGCTTAAACATTCATCAACATACCTTGCCTGGTGGTACCTTTGGTTGGTGAACAAAAT 1587  
DB 1653 ATGTTCAAATTTGAATCACTACCCCGCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712  
QY 1588 GAA 1590  
DB 1713 GAA 1715

RESULT 10  
US-08-331-384-2/c  
; Sequence 2, Application US/08331384  
; Patent No. 5856152  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Kelley, William M.  
; APPLICANT: Fisher, Krishna J.  
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods  
; TITLE OF INVENTION: of Use Thereof  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/331,384  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: UPNG1149USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4910 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; US-08-331-384-2

Query Match 5.3%; Score 107.8; DB 2; Length 4910;  
Best Local Similarity 53.4%; Pred. No. 8.9e-21;  
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 1168 GGCAGAGGAACTCTGTTTATTCATGGACCGCCAGCAGCAGGCAAAATCTATTATGCA 1227  
DB 1544 GGCAGAGGAACTCTGTTTATTCATGGACCGCCAGCAGCAGGCAAAATCTATTATGCA 1485  
QY 1228 CAAAGCCATAGCACAAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCAAATGTAACATTT 1287  
DB 1484 GAGGCCATAGCCCACTGTGCGCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1425  
QY 1288 CCATTAAATGACTGTACCAAGAACTTTGATTGGGTAGAAAGCTGTTAACTTTGGA 1347  
DB 1424 CCCTTCAACGACTGTGTCACAGATGTTGATCTGTTGGGAGGAGGAGATGACCGCC 1365  
QY 1348 CAGCAAGTAAACCAAGTTAAAGCCATTTGCTCTGCTCAAGCTATTTCGATTTGATCAAAA 1407  
DB 1364 AAGTCTGTGAGTGGCGCAAGCAATTTCTCGAGGAAGCAAGTGGCGGTGGACCAAGAA 1305  
QY 1408 GGAAGGAGCAGAAACAGATTGAACCAACACAGTCTATCATGACCAACAAATGAGACATT 1467  
DB 1304 TGCAAGTCTCGGCCAGATAGACCCGACTCCCGTGTATGTCACCTCCACCAACATG 1245  
QY 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527  
DB 1244 TGCGCGTGTATTCAGGGAACTCAACGACCTTCGAACACAGCAGCGTTGCAAGCCGG 1185  
QY 1528 ATGCTTAAACATTCATCAACATACCTTGCCTGGTGGTACCTTTGGTTGGTGAACAAAAT 1587  
DB 1184 ATGTTCAAATTTGAATCACTACCCCGCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1125  
QY 1588 GAA 1590  
DB 1124 GAA 1122

RESULT 11  
US-08-836-087-2/c  
; Sequence 2, Application US/08836087  
; Patent No. 5871982  
; GENERAL INFORMATION:  
; APPLICANT: Trustees of University of Pennsylvania  
; APPLICANT: Wilson, James M.  
; APPLICANT: Kelley, William M.  
; APPLICANT: Fisher, Krishna J.  
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
; TITLE OF INVENTION: Methods of Use Thereof  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, PO Box 457





Db 3539 ATGTTCAAAATTTGAACATCACCCCGCTCTGGATCATGACTTTTGGGAAGGTACCAAGCAG 3480  
QY 1588 GAA 1590  
Db 3479 GAA 3477

RESULT 15  
US-09-770-315-3  
; Sequence 3, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 7557  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-3

Query Match 5.3%; Score 107.8; DB 4; Length 7557;  
Best Local Similarity 53.4%; Pred. No. 1.1e-20;  
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 1168 GGCAGAAAGAAATACTGTTTATTTTCATGGACAGCCAGCAGCGCAATCTATTTATGCA 1227  
Db 1321 GGCAGAGGAACACCATCTGGCTGTTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 1380  
QY 1228 CAAGCCATAGCACAGCAGTTGGCAATGTTGGTTGCTATATATGAGCCCAATGTAAACTTT 1287  
Db 1381 GAGGCCATAGCCACACTGTCGCTTCTAGGGTGCGTAACTGGACCAATGAGAACTTT 1440  
QY 1288 CCATTTAATGACTGTACCAACAGAACTTGAATTTGGGTAGAAGAAGCTGGTAACTTTGGA 1347  
Db 1441 CCCTTCAACGACTGTGCGACAAGATGGTGTCTGGTGGGAGGAGGGAAGATGACCGCC 1500  
QY 1348 CAGCAGTAAACGAGTTTAAAGCAATTTGCTCTGGTCAAGCTATTCGCATTGATCAAAA 1407  
Db 1501 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGAGGAAGCAAGGTGCGCGTGGACCAAGAA 1560  
QY 1408 GGAAGGCGAGCAAAACAGATTGAACCAACACCAAGTATCATGACCAAAATGAGAACATT 1467  
Db 1561 TGAAGTCCTCGGCCCAAGATAGACCCGACTCCCGTGATGTCATCTCAACACCAACATG 1620  
QY 1468 ACAGTGTGTAGAAATAGGCTCGGAAGAAAGAACAGAACACACTCAACCAATCAGAGACAGA 1527  
Db 1621 TGCGCCGTGATTGACGGGAACCTCAACGACCTTCGAACACAGCAGCGCTTGAAGACCGG 1680  
QY 1528 ATGCTTAACATTATCTAACACATACCTTCGCTGGTGACTTTGGTTTGGTTGACAAAAT 1587  
Db 1681 ATGTTCAAATTTGAACCTCACCCCGCGTCTGGATCATGACTTTGGGAAGGTACCAAGCAG 1740  
QY 1588 GAA 1590  
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:06:01  
Job time : 169.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds  
(without alignments)

10821.780 Million cell updates/sec

Title: US-10-069-056-16

Perfect score: 2019

Sequence: 1 atggctggaagtcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1771	87.7	5121	18	US-10-647-111A-1
2	941.4	46.6	5075	17	US-10-361-002-3
3	941.4	46.6	5075	17	US-10-361-004-3
4	307.4	15.2	374	18	US-10-466-894-1276
5	285.8	14.2	421	18	US-10-466-894-430
6	284.8	14.1	423	18	US-10-466-894-375
7	284.8	14.1	464	18	US-10-466-894-370
8	284.8	14.1	473	18	US-10-466-894-367
9	284.8	14.1	486	18	US-10-466-894-358
10	284.8	14.1	491	18	US-10-466-894-428
11	284.2	14.1	420	18	US-10-466-894-438
12	283.8	14.1	343	18	US-10-466-894-388

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, App
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	117	5.8	1386	9	US-09-792-630-32	Sequence 32, Appl
40	117	5.8	1386	10	US-09-953-351-32	Sequence 32, Appl
41	117	5.8	1386	13	US-10-080-376-32	Sequence 32, Appl
42	117	5.8	1386	14	US-10-082-671-38	Sequence 38, Appl
43	117	5.8	1386	14	US-10-097-100-32	Sequence 32, Appl
44	115.4	5.7	1386	15	US-10-023-208-32	Sequence 32, Appl
45			1884	9	US-09-792-630-28	Sequence 28, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-647-111A-1  
; Sequence 1, Application US/10647111A  
; Publication No. US20040209240A1  
; GENERAL INFORMATION:  
; APPLICANT: IGGO, RICHARD  
; APPLICANT: MALERBA, MADDALENA  
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES  
; FILE REFERENCE: 604-691  
; CURRENT APPLICATION NUMBER: US/10/647,111A  
; CURRENT FILING DATE: 2003-08-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 5121  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1  
; OTHER INFORMATION: with promoter P4 and left hairpin from WMV (pph1)  
US-10-647-111A-1

Query Match	87.7%	Score	1771	DB	18	Length	5121
Best Local Similarity	92.3%	Pred. No.	0				
Matches	1864	Conservative	0	Mismatches	155	Indels	0
				Gaps	0		
Qy	1	ATGGCTGGAAATCCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
Db	282	ATGGCTGGAAATCCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
Qy	61	AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGGAATGGAATA	120				
Db	342	AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGGAATGGAATA	401				
Qy	121	GATATCGATGGAATAGTTACAAAAAGAGCTCGAGGAGCGAGCTGAAATCTTTACAA	180				

[illegible]

1482	CCAGCAGCACGGCAAATCTATTATTGCACAAGCCATTAGCAACGAGTTGGTAAATGTTGGT	1541
Qy	TGCTATAATGCGACCAAATGTFAAATCTTTCCAATTAAATGAATGTAACCAAGAAGACTTGATTT	1320
Dd	TGTTACAAATGCTGCCAATGTGAATCTTTCCAATTAAATGACTGTATACCAAAAAAATCTTGATTT	1601
Qy	TGGGTAGAGAAGCTGTGGTAACTTTTGGNACAGCAAGTAAAACCAAGTTTAAAGCCATTTGCTCT	1380
Dd	TGGGTGGAGAAGCTGTGTAACTTTTGGCCACGAAGTAAACCAATTCAAAGCTATTTTGTTCT	1661
Qy	GGTCAAGCTATTTCGCATTGTGATCAAAAGGAAAAAGGAGCAACAGATTTGAACCAACACCA	1440
Dd	GCCCAAACCATATGCAATTTGATCAAAAGGAAAGGAGGAGCAACAGATTTGAACCAACACCA	1721
Qy	GTCAATCATGACCAAAATATGAGAAATTTACAGTGGTCAGAAATAGCGTCGGAAGAAAGACCA	1500
Dd	GTTATTATGACCACCAACGAGAAATTTACCGTGGTTAGAAATAGCGTGTGAGGAAAGACCA	1781
Qy	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTCATCAACACATACCTTGCST	1560
Dd	GTTATTATGACCACCAACGAGAAATTTACCGTGGTTAGAAATAGCGTGTGAGGAAAGACCA	1841
Qy	GGTGACTTTCGGTTTGGTTGTGACAAAATATGAATGGCCCCATGATTTGTGCTTGGTTGGTAAAG	1620
Dd	GTTATTATGACCACCAACGAGAAATTTACCGTGGTTAGAAATAGCGTGTGAGGAAAGACCA	1901
Qy	AATGGTTTACCAATCTACCATGCGNAGCTACTGTGCTHAATGGGGCAAAAGTTCCTGATTTGG	1680
Dd	AATGGTTTACCAATCTACCATGCGGCTTGTACTGTGCTAAATGGGGCAAAAGTTCCTGATTTGG	1961
Qy	TCGAAAATCTGGCGGAGCCAAAGGTGCCAACTCCCTATAATTTACTTAGTTCGGCACGC	1740
Dd	TCAGAGACTTGGCGGAGCCGAAAGCTAGACACTCTCTATAATTTCCGTTAGGTTTCAATCGGC	2021
Qy	TCACCAATTCACGACACCCGAAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTTCCAATT	1800
Dd	TCACCAATCTGACTCCGCGAAGTACGCCCTCTCAGCCAAAATACGCTCTTACTTCCAATT	2081
Qy	GCATCGAATCTCGAGGACCTTGGCTTTAGAGCCTTTGGAGCACACCAANAATCTCCTGTTGGC	1860
Dd	GCATCGGACCTTGGCGACCTTAGCTCTAGAGCCTTTGGAGCACACCAANAATCTCCTGTTGGC	2141
Qy	GGCACTCGAGAAACCCAGAACACACTGGGGAGCTGGTTTCCAAAGCCTGCCAAGATGGTCAA	1920
Dd	GGCACTCGAGAACCCAGAACACACTGGGGAGCTGGTTTCCAAAGCCTGCCAAGATGGTCAA	2201
Qy	CTGAGCCCACCTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTTCGGGAACCG	1980
Dd	CGGAGCCCAACCTTGGTCCGAGATCGAGGCGGATTTGAGAGCGTGTCTTCAGTCAAGAACAG	2261
Qy	TTGAAGAAAGACTTCAGGAGCCCTGAACTTGGACTAA	2019
Dd	TTGAGAGCGCACTTCAACGAGGAGCTGACCTTGGACTAA	2300

```

RESULT 2
US-10-361-002-3
; Sequence 3, Application US/10361002
; Publication No. US20040170954A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gilmeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Armistead, David
; TITLE OF INVENTION: Pathogen Inactivation Assay
; FILE REFERENCE: CT-0043
; CURRENT APPLICATION NUMBER: US/10/361,002
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 5075

```



TYPE: DNA  
ORGANISM: Porcine parvovirus  
US-10-361-002-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;  
Best Local Similarity 70.4%; Pred. No. 5.7e-237;  
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

QY 4 GCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAAAGT 63  
Db 298 GCGGGAAACACTTACTCGGAAGAGGTACTAAAAAGCTACCAACTGGCTTCAAGATAATGCT 357  
QY 64 AACCGAGAGTGTCTCATTTGTTTAAAAAGTAAATGTTCAACTGAATGAAAAAGAT 123  
Db 358 CAAAAAGAGACATCTCTTATGTTTAAAAAGCAAAAAAGTCAATCTAAATGAAAAAGAA 417  
QY 124 ATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTCAACAGA 183  
Db 418 ATTGCTTGGATTAACCTACACAAGATACACAGATGCGGAATGATAAACCTTACAAGA 477  
QY 184 GGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGATGAA 243  
Db 478 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAATCGACAGC 537  
QY 244 ATGACCAAAAAGCAAGTATTCATTTTGTCTTTGTTTAAAAAGTATTTTGAAGTG 303  
Db 538 CTCACAAAACGGCAAGTACTGATTTTGTGACTCTCTTTTAAAAAGTGTCTCTTTGAAGGT 597  
QY 304 CTTTACACAAAAGATATTTCTGCTGTGATTTAAATTTGTTTCTGCAACATGAATGGGA 363  
Db 598 ATATTGCAAAAGAACCTAGTCCAGTGACTGCTACTGTTTCAACAGATGAACATGGT 657  
QY 364 AAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCTCAA 423  
Db 658 CAAGTACTGGCTATCACTGCCATGTACTAGTGGTGGAAAAGGCTTACACAAAGCAATG 717  
QY 424 GGAATATGGTGAAGAGGCACTAATATTTTACTGGAGCAGATGGTTGGTAACAGCCTGT 483  
Db 718 GGAATATGGTTCAGAAAAACAAATTTAAACAAATTTATGAGTAGATGGTTAATATGCAATGC 777  
QY 484 AATGTCAACTACACAGCTGAAGAAATTTAAACTAAGAGAAATAGCAGAAACAATAG 543  
Db 778 AAGTACTCTACACAGTTGAAGAAATTAATTTAAGGAATTTAGCAGAGATGGTGAG 837  
QY 544 TGGTTTACTCTACTTACTTATAGCATATAGCAACCAAAAAAGCACTATACCAAGTGTGT 603  
Db 838 TGGGTATCGCTACTAACCCTACACTCACAAACAATTAATAACAATATACAAAAATGACT 897  
QY 604 CTTTTTGGAAAAATGATGCTTACTATATTTTAACTAAAAAGAAATAAGCACTAGTCCA 663  
Db 898 CATTITGGAAATATGATTTGCTTACTACTTCTTAATAAAAAAGAAAGACAACT----- 951  
QY 664 CCAAGAGCGGAGCTATTTCTTACAGTGAATCTGGCTGGGAAACTTAACCTTTTAAAA 723  
Db 952 GAAAGAGAGATGGATATTTCTCAGCTCAGATTTCTGGCTTCAATGCAAAATTTCTTAAAA 1011  
QY 724 GAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCAGAAAAACGGTT 783  
Db 1012 GAAGGCGAGAGACACTTAGTCACTACCTATTTTACTGAAGCAATTAACCTGAACTGTG 1071  
QY 784 GAAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGCAAGAAATCAAACTAAAAAGAA 843  
Db 1072 GAAACCAAGCTTACTCAGCTCAGGAAGCCAAAAGAGGCGAGAAATACAAACAAAAAGAA 1131  
QY 844 GTTCTATTAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAGGAC 903  
Db 1132 GTAAGCATAAAAATGCAATAAAGAGCTTGGTTTAAATAAAGATGTACTAGCATAGAGAC 1191  
QY 904 TGGATGATGTCAGCAGACAGTATCATTTGAATGATGCTCAACCCAGGTGAGAAAAAC 963  
Db 1192 TGGATGATGACAGATCAGACGTTTATATGAATATGCTCTCAACCGGAGGAGAAAT 1251  
QY 964 CTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCATTT 1023

## RESULT 3

US-10-361-004-3

; Sequence 3, Application US/10361004

; Publication No. US20040170981A1

; GENERAL INFORMATION:

; APPLICANT: Clearant, Inc.

; APPLICANT: McKenney, Keith

; APPLICANT: Gillmeister, Lidja

; APPLICANT: Marlowe, Kristina

; APPLICANT: Armistead, David

; TITLE OF INVENTION: Real-Time Polymerase Chain Reaction Using Large Target Amplicons

; FILE REFERENCE: CI-0042

; CURRENT APPLICATION NUMBER: US/10/361,004

Db 1252 TTAATCAAANAATACACTAGAAATACAACTCTTACTCTAGCAAGACAAAAACAGCATAT 1311  
QY 1024 GACTTAATTTTAAAAAGCTGAAACCAAGCAAACTAAACCAACTTTTCACTGCGCTGACACA 1083  
Db 1312 GACTTAATTTTAAAAAGGCAAAACCAAGCATGCTACCAACATTTAATAATTAGCAATACA 1371  
QY 1084 AGNACCTCAGCAATTTTGTCTTTTCAATGCTGGAACATGTTAAAGTTTGGCAATGCTATT 1143  
Db 1372 AGNACATGTAATAATTTTCAAGCATGCAATTTGGAACCTACATTTAAAGTCTGCGCATATA 1431  
QY 1144 TGTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACTGTTTTTATTTTCAATGACACAGCC 1203  
Db 1432 ACTTGTTGTTAAACAGACAGAGGAAAGAAAGAAATACAACTCTTNTTNTTCAATGCGCAGCA 1491  
QY 1204 AGCACAGGCAAACTTATTATTGACAGGCAATAGCAAGCAAGCAAGTGTGGCAATGTTGTTGC 1263  
Db 1492 TCAACAGGAAAAAGTATAATTTGCTCAACATTTGCAAACTTAGTTGGTAATGTTGGTTGC 1551  
QY 1264 TATAATGACGCAATGTTAAACCTTTCCATTTAATGACTGTACCAACAGAACTTTGATTTGG 1323  
Db 1552 TACAATGACGCAATGTTAACTTTTCCATTTAATGACTGTACAAATAAAACTTTAATATGG 1611  
QY 1324 GTAGAGAGAGCTGTTAACTTTGGACAGCAAGTAAACCAAGTTTAAAGCATTTCCTCTGCT 1383  
Db 1612 ATTGAAGAGAGCAGGAACTTCTTAAACCAAGTAAACCAATTTCAAGCCATATGTTCAAGT 1671  
QY 1384 CAAGCTATTTCGCATTTGATCAAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACACCAAGTC 1443  
Db 1672 CAAACAAATTTAGATTTGACCCAAAAGGTTAAAGGAGCAACAAATTTGAACCACTCTCTGA 1731  
QY 1444 ATCATGACCAAAATGAGAACATTTACAGTGGTTCAGAAATAGGCTGCGAAAGAAAGACAGAA 1503  
Db 1732 ATAATGACTACAAATGAGACATACTAAAGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1791  
QY 1504 CACACTCAACCAATCAGACAGCAAGTCTTAACTTCACTTAACTTCACTTAACTTCACTTAACT 1563  
Db 1792 CATACAAACCAATTAAGACAGCAAGTGTAAACATAAACCTAACCAACCAACCACTGCGCAGT 1851  
QY 1564 GACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTTGTCTTGGTTGGTAAAGAAAT 1623  
Db 1852 GATTTTGGATTTTAGAAGAACTGAATGGCCATTAATATGCTTGGTTGGTAAAGAA 1911  
QY 1624 GGTACCAATCTACATGCGCAAGTACTGTGTTAAATGGGGCAAGTTCCTGATTTGCTCA 1683  
Db 1912 GGTACCAAGCAACAAATGCTAGCTATATGCTATATGCTATATGCTATATGCTATATGCTAT 1971  
QY 1684 GAAACTGGCGGAGCCAAAGTGGCCAACTCTTAAATTTTAACTTAACTTAACTTAACTTAACT 1743  
Db 1972 GAAATATGGAGGAGCCAAAAATGCAAAACCCCAATTAATAACCA---ACAGACTCTCAG 2028  
QY 1744 CCATTCAGCACACCGAAAAAGTAGCCCTCTCAGCCAGAACTATATGCACTTAACCTCACTTGCA 1803  
Db 2029 ATTTCACATCAGTGAATACTTGGCAGCGGCAACAACTACGCGCACTCCAAATACAG 2088  
QY 1804 TCGGATCTCGA---GGACCTGGCTTTAGAGCTTTGGAGCACACCAAAATACTCC 1853  
Db 2089 GAGGACCTGGATTTAGCTTTAGCTTTGAGCCGTGGAGCGGAGCCCAACACACC 2141

## RESULT 3

US-10-361-004-3

; Sequence 3, Application US/10361004

; Publication No. US20040170981A1

; GENERAL INFORMATION:

; APPLICANT: Clearant, Inc.

; APPLICANT: McKenney, Keith

; APPLICANT: Gillmeister, Lidja

; APPLICANT: Marlowe, Kristina

; APPLICANT: Armistead, David

; TITLE OF INVENTION: Real-Time Polymerase Chain Reaction Using Large Target Amplicons

; FILE REFERENCE: CI-0042

; CURRENT APPLICATION NUMBER: US/10/361,004

; CURRENT FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 5075  
; TYPE: DNA  
; ORGANISM: Porcine parvovirus  
US-10-361-004-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;

Best Local Similarity 70.4%; Pred. No. 5.7e-237; Indels 12; Gaps 3;  
Matches 1305; Conservative 0; Mismatches 536;

```
QY 4 GCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAAGT 63
DB 298 GCGGGAACACATTACTCGGAAGAGGTACTAAAGAGCTACCACTGGCTTCAAGATAATGCT 357
QY 64 AACGAGAGTGTCTTCATTTGTTTAAATGAAATGTTCAACTGTAATGGAAGAT 123
DB 358 CAAAAAGAGCATCTCTTATGTTTAAACACAAAAAGTCAATCTAAATGGAAAAAGAA 417
QY 124 ATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGCGAGCTGAAATCTTTACAAAGA 183
DB 418 ATTGCTTGGAAATACCTACAAAGAGATACACAGATGCGGAATGATTAACCTTACAAAGA 477
QY 184 GGAGCGGAATCTTGGGACCAAGCGAGGACATGGAATGGGAAACCCAGTGGATGAA 243
DB 478 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAATCGACAGC 537
QY 244 ATGACCAAAAGCAAGTATTCATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 303
DB 538 CTCACAAAACGGCAAGTACTGATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 597
QY 304 CTTAAACACAAAGATATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 363
DB 598 ATATTGCAAAAGAACCTTAAGTCCAAAGTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 657
QY 364 AAAGACCAAGGCTGGCACTGCCATGTAATTTGAGGAAAGGACTTTAGTCAAGCTCAA 423
DB 658 CAAGATATGCTATCACTGCCATGTAATTTGAGGAAAGGCTTTACAAACAGCAATG 717
QY 424 GGAATTTGGTGAAGAGGCACTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 483
DB 718 GGAATTTGGTTCGAAAGCACTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 777
QY 484 AATGTGCACTAACACAGCTGAAAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 543
DB 778 AAGTACCTTAACACAGTTGAAAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 837
QY 544 TGGGTTACTCTTACTTATAGCATAAGCAACCAAAAGAGCTATACCAAGTGTGTT 603
DB 838 TGGGTTACTCTTACTTATAGCATAAGCAACCAAAAGAGCTATACCAAAAGTGTGTT 897
QY 604 CTTTTTGGAAACATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 663
DB 898 CATTITGGAAATATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 951
QY 664 CCAAGAGACGGAGCTATTTCTTAGCAGTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 723
DB 952 GAAAGAGAGCATGGATATATCTCAGTCAAGTCTGAGTTCTGAGTTCTGAGTTCTGAGTTCTGAG 1011
QY 724 GAAGGCGAGCGCCATCTAGTGAGCAACTATACACTGATGACATGCGGCGAGAAACGGTT 783
DB 1012 GAAGGCGAGAGACATTAGTCACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1071
QY 784 GAAACCAAGTAACTAGTGGGAGGAACTAAGCGGCGGAGATTTCAACTTAAAGAA 843
DB 1072 GAAACCAAGTAACTAGTGGGAGGAACTAAGCGGCGGAGATTTCAACTTAAAGAA 1131
QY 844 GTTCTTATTAATACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCACCAGAGGAC 903
DB 1132 GTAAGCAATAAATGCACAATAAGAGACTTGGTTAATAAAGATGTACTAGCATAGAAGAC 1191
```

## RESULT 4

US-10-466-894-1276  
; Sequence 1276, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Anson, Robert  
; APPLICANT: Tuijnder, Marius

```
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-1276

Query Match      15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 2.3e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 174 TTTACACGAGGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACAC 233
Db 1 TTTACACGAGGCGGCGAGACCCTTGGGACCAAGCGAGGACATGGAATGGGAGGCG 60

QY 234 AGTGGATGAATGACCAAAAGCAAGTATCTTTTGGATCTTTGGTTAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATCTTTTGGATCTTTGGTTAAAGAGTGT 120

QY 294 ATTGGAAGTCTTAACCAAGAAATATATCTCTGGTGATGTTAAATGTTGCAACA 353
Db 121 GTTTGAAGTCTCAGCAAAAGCAATAGCTCTCTAGTAATGTTTGGTTGCGAGCA 180

QY 354 TGAATGGGAAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAAGACCAAGCTGGCACTGTCTGTCTGATTTGGAGCAAGGACTTAG 240

QY 414 TCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGGT 473
Db 241 TCAACCTCAAGGAAATGGTGGAGAGCAAGTAAATGTTTACTGGAGTANATGGTTGN 300

QY 474 AACAGCTGTAATGTGCAACTTAACCAAGCTGGAAGAAATTAAGTAAAGAAATGACGA 533
Db 301 GACTGCTGNAATGTNCAACTTAACCAAGCTGGAAGAAATTAAGTAAAGAAATGACGA 360

QY 534 AGACAATGAGTGG 546
Db 361 GGACAGTGANNNG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-375
```

```
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 276, 405, 417, 419, 421
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-430

Query Match      14.2%; Score 285.8; DB 18; Length 421;
Best Local Similarity 87.6%; Pred. No. 1.2e-64;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCA 232
Db 59 CTTTACACGAGGCGGCGAGACCCTTGGGACCAAGCGAGGACATGGAATGGGAGGCG 118

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATCTTTTGGATCTTTGGTTAAAAATGTT 292
Db 119 CAGTGGATGACATGACCAAAAGCAAGTATCTTTTGGATCTTTGGTTAAAGAGTGT 178

QY 293 TATTTGAAGTCTTAACCAAAAGAAATATATTTCTCGTGAATGTAATGTTGTGCAAC 352
Db 179 TGTTCGAAGTCTCAGCAAAAGCAATAGCTCTCTAGTAATGTTTACTTGGTTCGCGAGC 238

QY 353 ATGAATGGGAAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACTTTA 412
Db 239 ATGAATGGGAAAAGACCAAGCTGGCACTGTCTGATGTTGGAGGCAAGGACTTTA 298

QY 413 GTCAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGG 472
Db 299 GTCAGCTCAAGGAAATGGTGAGAGGCAAGTAAATGTTTACTGGAGTATGATGGTTGG 358

QY 473 TAACAGCTGTAATGTGCAACTTAACCAAGCTGGAAGAAATTAAGTAAAGAAAT 527
Db 359 TGACTGCTGTAATGTTCAACTTAACCAAGCTGGAAGAAATTAAGTAAAGAAAT 413

RESULT 6
US-10-466-894-375
; Sequence 375, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-375
```



Db 372 TGACTGCTGTAATGTTCAACTAACACAGCTGAAAAA 415  
|||||

## RESULT 9

US-10-466-894-358  
; Sequence 358, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 358  
; LENGTH: 486  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 413, 428, 435, 448, 457, 462  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;  
Best Local Similarity 89.2%; Pred. No. 2.4e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 173 CTTTACAAAGGAGCGGAACTACTTTGGACCAAAAGCAGGACATGGAATGGAAACCA 232  
Db 54 CTTTACAAAGGAGCGGAGACCACTTTGGACCAAAAGCAGGACATGGAATGGAGAGCG 113  
Qy 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGGTCTTTGTTTAAAAATGTT 292  
Db 114 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTCTTTGTTTAAAAATGTT 173  
Qy 293 TATTGAAGTCTTAAACAAAGAAATATTTCTGCTGATGTTAATTTGTTTGTGCAAC 352  
Db 174 TGTGTAAGTCTCAGCACAAAGAAATAGCTCTCTAGTAATTTTCTGTTGTTGTCGAGC 233  
Qy 353 ATGAATGGGAAAGCAAGGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412  
Db 234 ATGAATGGGAAAGCAAGGCTGGCAGCTGTCTGCTGATTTGGAGGCAAGACTTTA 293  
Qy 413 GTCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATGTTTCTGAGCAGATGTTGG 472  
Db 294 GTCAAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTGAGTGTGAGTGTGTTGG 353  
Qy 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAAGAAATTA 516  
Db 354 TGACTGCTGTATGTTCAACTAACACAGCTGAAAGAAATTA 397  
|||||

## RESULT 10

US-10-466-894-428  
; Sequence 428, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 428  
; LENGTH: 491  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 12, 421, 441  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;  
Best Local Similarity 89.2%; Pred. No. 2.4e-64;  
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 173 CTTTACAAAGGAGCGGAACTACTTTGGACCAAAAGCAGGACATGGAATGGAAACCA 232  
Db 62 CTTTACAAAGGAGCGGAGACCACTTTGGACCAAAAGCAGGACATGGAATGGAGAGCG 121  
Qy 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGGTCTTTGTTTAAAAATGTT 292  
Db 122 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTCTTTGTTTAAAAATGTT 181  
Qy 293 TATTGAAGTCTTAAACAAAGAAATATTTCTGCTGATGTTAATTTGTTTGTGCAAC 352  
Db 182 TGTGTAAGTCTCAGCACAAAGAAATAGCTCTCTAGTAATTTTCTGTTGTTGTCGAGC 241  
Qy 353 ATGAATGGGAAAGCAAGGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412  
Db 242 ATGAATGGGAAAGCAAGGCTGGCAGCTGTCTGCTGATTTGGAGGCAAGACTTTA 301  
Qy 413 GTCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATGTTTCTGAGCAGATGTTGG 472  
Db 302 GTCAAGCTCAAGGAAATGTTGAGAGGCAAGTAAATGTTGAGTGTGAGTGTGTTGG 361  
Qy 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAAGAAATTA 516  
Db 362 TGACTGCTGTATGTTCAACTAACACAGCTGAAAGAAATTA 405  
|||||

## RESULT 11

US-10-466-894-438  
; Sequence 438, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; CURRENT FILING DATE: 2003-07-23  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 438

```
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 418
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-438

Query Match      14.1%; Score 284.2; DB 18; Length 420;
Best Local Similarity 87.8%; Pred. No. 3.2e-64;
Matches 310; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 173 CTTTACAAACGAGGCGGAAACTACTTGGACCAAAAGCAGGACATGGAATGGGAAACCA 232
Db      |||
Qy 64 CTTTACAAACGAGGCGGAGCCACTTGGACCAAAAGCAGGACATGGAATGGGAGGCG 123
Db      |||
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGATTCTTTGGTTAAAAAATGTT 292
Db      |||
Qy 124 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGATTCTTTGGCTTAAGAAGTGT 183
Db      |||
Qy 293 TATTGGAAGTCTTAACACAAAGAAATATTTCTCTGCTGATGTTAATTTGGTTGTGCAAC 352
Db      |||
Qy 184 TGTTTGAAGTGTCTCAGCACAAAGAAACATAGCTCTTAGTAATGTTACTTTGGTTTGTGAGC 243
Db      |||
Qy 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTA 412
Db      |||
Qy 244 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTCATGTCTGCTGATTGGAGGCAAGGACTTTA 303
Db      |||
Qy 413 GTCAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTTGGAGCAGATGTTGG 472
Db      |||
Qy 304 GTCACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTTGGAGTAGATGTTGG 363
Db      |||
Qy 473 TACAGCTCTAATGTGCACTAACACACAGCTGAAAGAAATTAACCTAAGAGAA 525
Db      |||
Qy 364 TGACTGCTGTAAATGTTCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAA 416
Db      |||

RESULT 12
US-10-466-894-388
; Sequence 388, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijinder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 388
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-388

Query Match      14.1%; Score 283.8; DB 18; Length 343;
Best Local Similarity 89.2%; Pred. No. 3.6e-64;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 174 TTTTACAAACGAGGCGGAAACTACTTGGACCAAAAGCAGGACATGGAATGGGAAACCA 233
Db      |||
Qy 1 TTTTACAAACGAGGCGGAGACCACTTGGACCAAAAGCAGGACATGGAATGGGAGGCGC 60
Db      |||
Qy 234 AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGGATTCTTTGTTAAAAAATGTTT 293
Db      |||
Qy 61 AGTGGATGACATGACCAAAAGCAAGTATTATTTTGGATTCTTTGTTAAAGAAGTGT 120
Db      |||
Qy 294 ATTTGAAGTCTTAACACAAAGAAATATTTCTCTGCTGATGTTAATTTGGTTTGTGCAAC 353
Db      |||
Qy 121 GTTTGAAGTGTCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTTACTTTGGTTCTGCGAGCA 180
Db      |||
Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTAG 413
Db      |||
Qy 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTCATGTCTGATTGGAGGCAAGGACTTTAG 240
Db      |||
Qy 414 TCAAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTTGGAGCAGATGTTGGT 473
Db      |||
Qy 241 TCAACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTGTAGTAGATGTTGGT 300
Db      |||
Qy 474 AACAGCTGTAAATGTGCACTAACACAGCTGAAAGAAATTAAC 516
Db      |||

Query Match      14.1%; Score 283.8; DB 18; Length 343;
Best Local Similarity 89.2%; Pred. No. 3.6e-64;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 174 TTTTACAAACGAGGCGGAAACTACTTGGACCAAAAGCAGGACATGGAATGGGAAACCA 233
Db      |||
Qy 1 TTTTACAAACGAGGCGGAGACCACTTGGACCAAAAGCAGGACATGGAATGGGAGGCGC 60
Db      |||
Qy 234 AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGGATTCTTTGTTAAAAAATGTTT 293
Db      |||
Qy 61 AGTGGATGACATGACCAAAAGCAAGTATTATTTTGGATTCTTTGTTAAAGAAGTGT 120
Db      |||
Qy 294 ATTTGAAGTCTTAACACAAAGAAATATTTCTCTGCTGATGTTAATTTGGTTTGTGCAAC 353
Db      |||
Qy 121 GTTTGAAGTGTCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTTACTTTGGTTCTGCGAGCA 180
Db      |||
Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTAG 413
Db      |||
Qy 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTCATGTCTGATTGGAGGCAAGGACTTTAG 240
Db      |||
Qy 414 TCAAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTTGGAGCAGATGTTGGT 473
Db      |||
Qy 241 TCAACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTGTAGTAGATGTTGGT 300
Db      |||
Qy 474 AACAGCTGTAAATGTGCACTAACACAGCTGAAAGAAATTAAC 516
Db      |||
```

Db 301 GACTGCCTGTAATGTTCAACTTAACACGAGCTGAAAAA 343

## RESULT 14

US-10-466-894-433  
; Sequence 433, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent  
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; PRIOR FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 433  
; LENGTH: 403  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 218  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-433

Query Match 14.1%; Score 283.8; DB 18; Length 403;  
Best Local Similarity 89.0%; Pred. No. 4e-64;  
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGAAACCA 232  
Db 56 CTTTACACGAGGCGGAGACCACTTGGACCAAAAGCGAGGACATGGAATGGAGAGCG 115  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATGTT 292  
Db 116 CAGTGGATGATGACCAAAAGCAAGTATTTATTTTGTATCTTTGTTTAAAGAGTGT 175  
QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTGTGTATGTTAATTTGTTGCAAC 352  
Db 176 TGTTTGAAGTCTCAGCACAAAGACATAGCTCCTAGTAATGNTACTTGTGTCGAGC 235  
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412  
Db 236 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGTCTATGCTGATTTGGAGGCAAGACTTTA 295  
QY 413 GTCAAGCTCAAGGAAATGTTGAGAAAGGCAACTAAATGTTTACTTGAGCAGATGTTGG 472  
Db 296 GTCAACCTCAAGGAAATGTTGAGAAAGGCAAGTAAATGTTGACTGGAGTAGATGTTGG 355  
QY 473 TAACAGCTGTATGTGCAACTTAACACAGCTGAAGAATTTAA 516  
Db 356 TGACTGCCTGTATGTTCAACTTAACACAGCTGAAGAATTTAA 399

## RESULT 15

US-10-466-894-369  
; Sequence 369, Application US/10466894  
; Publication No. US20040241671A1  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Tuijnder, Marius  
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,  
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE  
; FILE REFERENCE: 10918-014-999  
; CURRENT APPLICATION NUMBER: US/10/466,894  
; PRIOR FILING DATE: 2003-07-23  
; PRIOR APPLICATION NUMBER: PCT/FR02/00273  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: FR01/00899  
; PRIOR FILING DATE: 2001-01-23  
; NUMBER OF SEQ ID NOS: 2270  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 369  
; LENGTH: 497  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 448,475  
; OTHER INFORMATION: n = A,T,C or G  
US-10-466-894-369

Query Match 14.0%; Score 283.2; DB 18; Length 497;  
Best Local Similarity 89.0%; Pred. No. 6.4e-64;  
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGAAACCA 232  
Db 68 CTTTACACGAGGCGGAGACCACTTGGACCAAAAGCGAGGACATGGAATGGAGAGCG 127  
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATGTT 292  
Db 128 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTATCTTTGTTTAAAGAGTGT 187  
QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTGTGTATGTTAATTTGTTGCAAC 352  
Db 188 TGTTTGAAGTCTCAGCACAAAGACATAGCTCCTAGTAATGTTTGTGTCAGC 247  
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412  
Db 248 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGTCTATGCTGATTTGGAGGCAAGACTTTA 307  
QY 413 GTCAAGCTCAAGGAAATGTTGAGAAAGGCAACTAAATGTTTACTTGAGCAGATGTTGG 472  
Db 308 GTCAACCTCAAGGAAATGTTGAGAAAGGCAAGTAAATGTTGACTGGAGTAGATGTTGG 367  
QY 473 TAACAGCTGTATGTGCAACTTAACACAGCTGAAGAATTTAA 516  
Db 368 TGACTGCCTGTATGTTCAACTTAACACAGCTGAAGAATTTAA 411

Search completed: January 23, 2005, 14:35:20  
Job time : 1079 secs

**This Page Blank (uspto)**